The Journal of Applied Psychology

Vol. V SEPTEMBER, 1921 No. 3

EDITORS

JAMES P. PORTER
Clark University

WILLIAM F. BOOK
Indiana University

And a Board of Co-operating Editors

CLARK UNIVERSITY
WORCESTER MASS.
FLORENCE CHANDLER, Publisher

Entered as second class matter June 6, 1917, at the post office at Worcester, Mass. under the Act of March 3, 1879.

Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917. Authorized April 26, 1921.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. M. Bouchard, Carnegie Institute of Technology; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Harvard University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, University of Wisconsin; Edward K. Strong, Jr., George Peabody College for Teachers; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Psycho-Educational Clinic, St. Louis; G. M. Whipple, University of Michigan; Helen Thompson Woolley, Vocation Bureau, Cincinnati.

The scope of the Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeal. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environmental conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

VOLUME V SEPTEMBER, 1921 NUMBER 3

TABLE OF CONTENTS

F. Kuhlmann. The Results of Repeated Mental Re-Examinations of 639 Feeble-Minded Over a Period of Ten Years 195

William Healy. Pictorial Completion Test II 225

Frank Watts. The Construction of Tests for the Discovery of Vocational Fitness 240

D. S. Herrick. A Comparison of Brahman and Panchama Children in South India with each other and with American Children by Means of the Goddard Form Board 253

M. L. Ream. A Statistical Method for Incomplete Order of Merit Ratings 261

Kwok Tsuen Yeung. The Intelligence of Chinese Children in San Francisco and Vicinity 267

M. A. Bills. Methods for the Selection of Comptometer Operators and Stenographers 275

Frances L. Gaw. Minor Studies in Educational Psychology 284

I. School Ratings and Moving Pictures.

II. Relations of Stanford Tests and Dearborn Maze Tests.

Communications and Notes. Industrial Psychology in Europe 287

Book Reviews 292

Communications, manuscripts, reviews, summaries, queries, books and off-prints for review should be addressed to Dr. James E. Porter, Clark University, Worcester, Mass. Contributors are requested to follow the conventional method of citation employed by this Journal. In every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published quarterly. The subscription price is $4.00 a year. Single numbers $1.25. Remittances and business communications should be addressed to Florence Chandler, Publisher, Clark University, (950 Main St.), Worcester, Mass.

THE
JOURNAL OF APPLIED PSYCHOLOGY

VOL. V SEPTEMBER, 1921 No. 3

THE RESULTS OF REPEATED MENTAL RE-EXAMINATIONS OF 639 FEEBLE-MINDED OVER A PERIOD OF TEN YEARS

By F. KUHLMANN

Early in 1912 the writer adopted the intelligence quotient method of grading intelligence. This was before Stern's monograph in which this method is proposed appeared.[1] Results from the examination of 1,300 feeble-minded of all ages quoted at that time seemed to indicate that this ratio of mental age to age remained roughly constant through the increasing ages of the feeble-minded. It was pointed out, however, that this constancy required that normal or average mental development progress at a uniform rate from year to year, as measured in absolute units of measurement, an assumption contrary to general observation and theory. In 1911 a ten-year program was begun of re-examining all inmates of the Minnesota School for Feeble-Minded at regular intervals of two years, excluding cases over twenty years of age, all epileptics, and others in whom some special trait interfered with getting a reliable mental age. The object of this study was to test out thoroughly the value of the I. Q. method of classification, and to determine the traits of the mental growth curve.

The examinations involve 639 cases from the grade of idiocy to nearly average normal intelligence, and from one to twenty years of age. Each case was examined from two to five times. About a third of the examinations were made by the writer. The others were made by Maud A. Merrill, Dr. Frances Lowell, Katharine B. Graves, and Rose Anderson, past and present research assistants. All had had ex-

[1] See Kuhlmann, F. "Degree of mental deficiency in children as expressed by the relation of age to mental age." Journ. Psycho-Asthenics, June, 1913.

tensive preparation in psychology and thorough training in the use of the tests. Variability of results due to different examiners was probably a negligible factor.

Three disturbing factors were encountered. The first concerned the cases examined. A case once examined did not always remain in the institution for future re-examinations. Some were absent temporarily at the time a re-examination was due. Others were removed permanently through death, discharge, or transference to other institutions. The second factor was the possible selective influence on grade of intelligence of the institution in admitting and discharging cases at different ages. It could not be assumed that cases admitted or discharged at different ages were of the same grade of intelligence. The third factor was the varying accuracy of the tests at different age-levels, and of the different revisions of the B.-S. scale that were used. The first examinations were made in 1910 with the original 1908 scale.[2] From 1912 to 1917 my 1912 revision of the tests was used.[3] Since the beginning of 1917 my last revision was used. The first two gave on the whole quite the same results, the 1912 revision being concerned chiefly with standardization of procedure and extending the scale downwards below the age of three.[4] The last revision corrected the general tendency of the earlier scales of giving too high mental ages towards the lower end, and too low mental ages towards the upper end, and increased the number of tests to eight for each age group.

The first and second of these disturbing factors could be met satisfactorily. The influence of the third could be eliminated only in part. This will be considered in connection with the statement and discussion of the data.

Average Yearly Increase in Mental Age

The raw data gives a varying number of cases whose first examinations came at different ages, from one to eighteen years. Except for the disturbing factors just mentioned, the average mental growth curve for these inmates of the institution could be at once determined by computing the average mental age at each age of all cases examined and re-examined.

[2] See Kuhlmann, F. "Binet and Simon's system for measuring the intelligence of children." Journ. Psycho-Asthenics, 1911.

[3] See Kuhlmann, F. "A revision of the Binet-Simon system for measuring the intelligence of children." Journ. Psycho-Asthenics, Monograph Supplements, Sept., 1912.

[4] See Kuhlmann, F. "Some results of examining a thousand public school children with a revision of the Binet-Simon tests." Journ. Psycho-Asthenics, March and June, 1914.

This procedure, as a matter of fact, gives a growth curve surprisingly close to what is obtained when the disturbing factors are eliminated. The number of cases at each age is large enough to give a fairly smooth curve in spite of the fact that at each age some of the previous cases have dropped out and other new ones have been added. The selective influence of the institution, as shown also by special methods of treating the results to determine it, is small, and negligible for most parts of the growth curve. The mental ages at each age, since the cases include all grades, include a considerable range, and thus partly eliminate or smooth out the varying errors in the mental ages due to inaccuracies in the scale at different levels.

But these relationships change when we attempt to determine the mental growth curves separately for the different grades of intelligence. The difficulty is increased by the fact that our definition of grade of intelligence in terms of the intelligence quotient involves us in assumptions the correctness of which constitutes our chief problem. We may attribute a certain range of I. Q.'s to a given grade of intelligence for any particular age, but to use the same range for all ages for this grade of intelligence assumes, of course, that the I. Q. remains constant for all ages.

Table I gives the average mental ages for five successive years of each group of cases whose first examinations occurred at the same age. This is given separately for the four grades of intelligence termed idiots, imbeciles, morons, and borderline cases. The manner of deriving these figures needs to be carefully noted in detail. It was done through the I. Q.'s, as that procedure had certain advantages over that of dealing directly with the mental ages. The first step was to correct the individual mental ages found for errors due to the general tendency of the scale to measure too high or too low at different levels. This was done by subtracting from each individual mental age the amount the scale was known to measure too high on the average at this point. There were practically no mental ages at levels high enough to need additions as corrections. These corrections could not be made safely for mental ages below four, and could therefore not be extended to low grade imbeciles, idiots, or to the younger morons and borderline cases. Our norms for these lower mental ages are not sufficiently well established for any scale to allow of any definite conclusion as to what minor degree of error they may still contain. The data used in making these corrections were the smoothed figures on the average mental ages found with my 1912 revision on 1,000 public

school children, referred to above, and unpublished data for my last revision.

The second step was to compute the I. Q.'s for the ages between two successive examinations by taking the average of the one preceding and the one following. It was then easier and more accurate to compute the average mental age at a given age from the average I. Q. than it would have been to work with the mental ages directly. The examinations did not, of course, occur at the exact ages of seven, eight, and so on, and to have averaged the mental ages of all cases classified as examined at seven, for example, might have been more or less misrepresentative, as the exact average age of these cases could vary from six and a half to seven and a half years. All age classifications were made on the basis of the nearest birthday. The error was eliminated by computing the average mental age from the average I. Q., using the ages as exactly seven, eight, and so on. Results of fourth and fifth examinations of a case were thrown out in order to put all cases on the same basis of a five year period with three actual examinations at two year intervals. Data for ages below seven are not given in this table because the mental ages could not be corrected for the general tendency to error in the scale, in the case of the morons, for ages below this.

The third step aimed to eliminate any selective influence of the institution there might have been on grades of intelligence at different ages, and also the difficulty arising from classifying the cases into the four grades of intelligence at different ages on the basis of the I. Q. This procedure was as follows. (a) Computing the average I. Q. for morons, for example, examined first at the age of seven, and again at nine and eleven. (b) Computing the average I. Q. at eight for those examined first at eight, and again at ten and twelve. (c) Eliminating from this second group of cases low or high grades cases so that the average I. Q. of the remaining cases was approximately the same as the computed average I. Q. at eight for the cases examined first at seven. If there was a general tendency of the I. Q. to change from seven to eight, this procedure insured having the same true grade of intelligence for the group examined first at eight as for the group examined first at seven. On account of the small number of cases at times for such a group, this equating of the I. Q.'s could be done only approximately in many instances. The remaining step was taken to make such exact equating unnecessary. (d) The average I. Q.'s at ten and twelve were

TABLE I

IDIOTS

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6		1.60	1.88	2.17	2.44	2.72								
4			1.88	2.17	2.48	2.53	2.54							
5				2.17	2.36	2.55	2.71	2.86						
11					2.43	2.65	2.88	3.04	3.19					
2						2.61	2.78	2.93	2.97	2.81				
3							2.73	2.91	3.08	3.04	3.00			
2								2.93	3.12	3.10	3.08	3.07		
3									3.09	3.04	2.99	2.87	2.75	
										3.00	3.03	2.97	2.75	

IMBECILES

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
8	2.80	3.17	3.53	3.91	4.28									
14		3.17	3.53	3.88	4.14	4.38								
16			3.53	3.77	3.99	4.28	4.44							
21				3.85	3.87	4.40	4.68	4.94						
24					4.07	4.21	4.30	4.38	4.44					
18						4.32	4.44	4.52	4.75	4.66				
25							4.46	4.71	4.94	4.84	4.73			
11								4.64	4.88	4.79	4.73	4.67		
12									4.75	4.74	4.74	4.76	4.78	
14										4.76	4.71	4.65	4.79	4.92
10											4.73	4.33	3.93	3.85
7												4.60	4.68	4.75
													4.55	4.51

MORONS

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	4.32	5.02	5.75	6.36	7.00									
10		5.02	5.64	6.26	6.96	7.57								
31			5.70	6.17	6.64	7.08	7.49							
24				6.26	6.65	7.00	7.44	7.94						
25					6.81	7.24	7.80	8.13	8.46					
21						7.22	7.55	7.96	8.40	8.28				
17							7.57	8.03	8.48	8.43	8.40			
24								8.01	8.63	8.77	8.73	8.80		
17									8.49	8.55	8.59	8.63	8.73	
17										8.49	8.51	8.58	8.67	8.73
14											8.56	8.41	8.39	8.35
13												8.61	8.41	8.25
													8.55	8.44

BORDERLINE

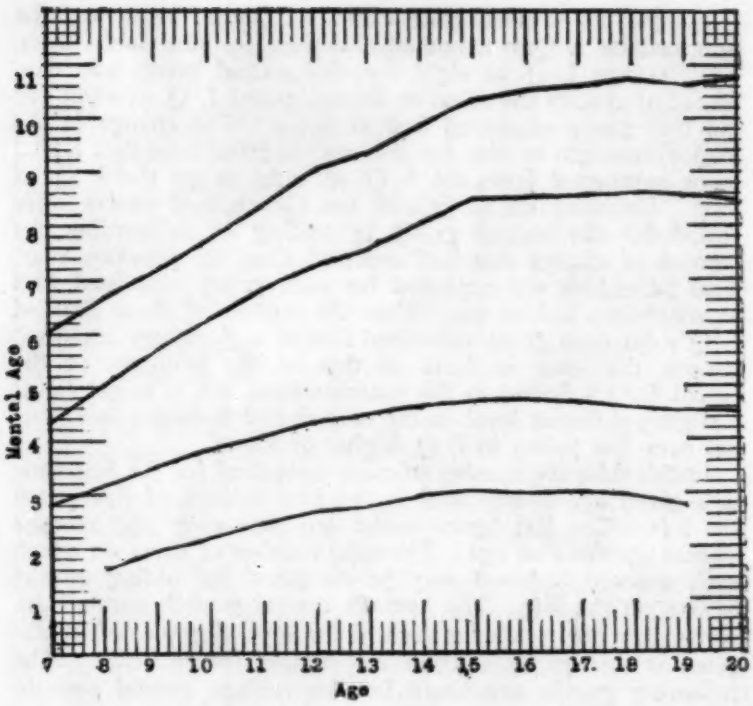
No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6	6.07	6.63	7.14	7.57	8.07									
9		6.63	7.25	7.95	8.71	9.47								
10			7.19	7.80	8.59	9.35	9.64							
10				7.77	8.10	8.69	9.31	10.05						
5					8.37	8.77	9.11	9.45	9.73					
5						9.07	9.46	9.83	10.53	10.53				
12							9.38	9.94	10.64	10.59	10.65			
10								9.82	10.62	10.64	10.76	10.77		
5									10.38	10.32	10.38	10.76	10.41	
4										10.52	10.63	10.65	10.80	11.03
5											10.61	10.70	10.79	10.98
8												10.78	10.46	10.14
													10.61	10.68

found for the selected cases examined first at eight, and the intermediate I. Q.'s computed as before for nine and eleven. The average I. Q. at eight for this second group was then placed at exactly the same as the computed I. Q. at eight for the first group examined first at seven. The change in the I. Q. from eight to nine for this second group was then added to or subtracted from the I. Q. at eight to get the I. Q. at nine. Likewise, the I. Q.'s at ten, eleven, and twelve were found for the second group by adding or subtracting the amount of change that had occurred since the previous year. This procedure was repeated for each group examined first at nine, ten, and so on. Thus the course of these derived I. Q.'s for each group examined first at a given age remained exactly the same in form as that of the averages of the actual I. Q.'s found in the examinations, but it might lie at a slightly different level, in the majority of instances in a level not over five points in I. Q. higher or lower.

In this table the number of cases examined for the first time at a given age is indicated in the first column of figures on the left. The last figure under any age gives the average mental age for that age. The total number of cases on which each average is based may be computed by adding up the figures on the left. The average mental growth curve from seven to twenty years for each of the several grades of intelligence is then given in these last figures for each age. The following graphs are based on the average mental ages in Table I.

The main features of these growth curves present but little that is not now pretty well known. They show that the feeble-minded do develop mentally, and at a rate in proportion to the grade of intelligence. This was still a disputed point at the time this study was begun. This rate of development decreases with age for all grades. The rate of this decrease cannot, of course, be shown with this data, as we do not know how much mental growth in terms of absolute units the mental year represents at different age levels. In the graphs each addition of a year in mental age is represented as equal to preceding additions of a year, but even so the curves decline with increasing age. Could they be plotted in terms of absolute units of growth they would decline very much more. This will be made more obvious below, in discussing the intelligence quotient.

Age of Cessation of Mental Age Increase. Space will not be taken here to review in detail previous studies and dis-



cussions on this matter. It will be remembered that Stern[5] assumed that the age of cessation of mental development changed markedly with grade of intelligence, the idiot grade ceasing developing relatively very soon after birth. Others since have accepted this assumption. Recently Doll[6] has claimed that his results on the re-examination of feeble-minded verified this assumption of earlier arrest for the lower grades. Grading his cases on the basis of the final mental ages they attain in reaching maturity, he summarizes as follows.[7] I

Final mental age....	1	2	3	4	5	6	7	8	9	10
Age of arrest.....	?	?	7	11	11	12	12	12	15	15

find myself unable to agree with Doll that his own results in this study will bear the interpretation given in these figures.

[5] See Stern, W. "The psychological methods of testing intelligence." Trans. by G. M. Whipple, 1914.

[6] See Doll, E. A. "The growth of intelligence." Princeton contributions to psychology, 1920.

[7] See page 76.

The Goddard revision of the B.-S. tests, which was used in this study, is not adequate for the determination of mental ages much below five, as it does not include standardized tests for ages below three. Doll's average growth curve for cases with a final mental age between 6 and 7 rises markedly beyond the age of 12. The same is true of his next curve, for cases with a final mental age between 7 and 8. His curve for cases with a final mental age between 9 and 10, and the next for cases with a final mental age between 10 and 11 show a still more obvious rise beyond the age of 15.[8] In fact, my own interpretation of Doll's results would be that they do not disagree seriously with my present results, based on a larger number of cases. In considering my own table and graphs, it should be noted that the different grades are defined in terms of the I. Q. at the age of seven,—eight for the idiot grade. At this age, cases with I. Q.'s from 0 to 24 were classed as idiots, cases with I. Q.'s from 25 to 49 as imbeciles, 50 to 74 as morons, and over 74 as borderline. The upper and lower limits of these ranges of I. Q.'s for these different grades then decreased with age, as explained above. It is seen that the average growth curves extending from the age of seven to twenty all rise to the age of fifteen at least. They indicate that idiots, as here defined, develop to the age of fifteen, inclusive, imbeciles to fifteen or sixteen, morons to about seventeen, and borderline cases to about eighteen. In other words, all grades develop much longer than Stern and others have supposed, and the difference for different grades is not nearly as great as has been assumed. These facts are shown even more strikingly in the I. Q. curves below than in the present mental age data. Since the mental age of cases of fifteen years or over was always divided by fifteen to get the I. Q., the I. Q. rises beyond this age, when the mental age continues to increase.

Frequency of Yearly Gain and Loss in Mental Age. In considering this question the results for all grades of cases will be combined. It will not be necessary to take into account the three disturbing factors noted above. Not correcting the mental ages for errors in the scale will also enable us to include all the results for mental ages below seven. Since the mental age increase for the two-year interval between two successive examinations is nearly always much less than two years, and since any two successive mental ages will both be affected in the same direction through the error in the scale, the amount of loss or gain in mental age between two

[8] See page 75.

successive examinations is hardly ever seriously affected by the error the scale made. The gain or loss in mental age for each case was computed in terms of the average number of months per year. That is, for each case, the average yearly change in months between the first and second examinations was computed, and the same for the two-year interval between the second and third examination, and so on. These two or three yearly averages for a case were then again averaged, and this final average was taken as his average rate of gain or loss. The average period covered for a case was about five years. This procedure smoothes out the effect of errors in any given mental age due to not getting the best efforts from a case during any particular examination, or to other unknown factors. A number of quite questionable results were eliminated, reducing the total to 600 cases. The following table gives the number of cases that gained or lost 1, 2, 3, etc., months a year in mental age. It includes all grades.

TABLE II

	Gain																				Loss						
Mos.....	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7
No.....	1	1	1	2	3	3	3	5	10	12	12	18	22	25	32	30	37	61	81	78	66	37	32	16	6	4	1	1
%	4.8											68									11	16						

For the percentages in the last line in the table the cases are divided into four classes; first, those that gain at a rate of twelve months or more a year, that is, at a rate equal to or greater than that of the average normal rate; second, those that gain one to eleven months a year; third, those that are practically standing still, gaining or losing less than a month a year; and fourth, those that actually lose in mental age, the deteriorating class. It is probable that these four classes represent more or less distinct types, the difference between which might perhaps be accounted for if we had a detailed knowledge of their history, heredity, health records, and physical conditions. This data is at present not available for these cases. It is particularly significant to find 4.8 per cent who gain at a rate greater than that of the average normal rate, and 16 per cent that deteriorate. There are undoubtedly some cases in the first class whose rapid gain is due to a too low and erroneous mental age for the first examination.

But the majority of them cannot be explained in that way, as the large yearly gain is shown throughout the course of three or more examinations. Barring the erroneous results, all cases of this class are improving in *rate* of development, as may be true, of course, of some of the 68 per cent in the next class. A very small number in the improving class continue this greater than average normal rate long enough to make up their deficiency entirely. They attain a final I. Q. of 1.00 or more.

For the same reason that most of these cases showing unusual gain cannot be explained on the grounds of an erroneous mental age for the first examination, the majority of cases comprising the 16 per cent that lose in mental age cannot be explained on the grounds of too low a mental age for the last examination. The tendency to deteriorate shows a marked relationship to grade, as was to be expected. Attributing the I. Q. ranges given above to the four grades, irrespective of age, shows the following percentages of this deteriorating class belonging to these grades.

Idiots	35	per cent
Imbeciles	23	"
Morons	12	"
Borderline	2	"

It should be noted that it is, of course, easier for the lower grade cases to show actual loss in mental age through an accidentally too low a score in the last examination than it is for the higher grade cases. The small accidental error of a few months may more easily exceed the real gain of a low grade case than it can the larger real gain of the higher grade case.

The tendency to deteriorate is also definitely related to age. The loss in mental age occurs more frequently with the older cases. In the following table the age figures in the first column indicate the ages at which the first examination was given. The figures in the next three columns give the number of cases under the respective classes, and the last column gives the percentage of cases at each age that lose in mental age.

These figures indicate a rather abrupt increase in the percentage losing in mental age at the age of eleven, which increase on the whole rises up to the final age of twenty, beyond which no examinations were made. It was expected that this increase with age in the percentage losing would come more from the lower than from the higher grades. When, however, the idiots and imbeciles, on the one hand, and the morons and borderline cases, on the other hand, are

TABLE III

Age	Number Gaining	Number Stationary	Number Losing	Per cent Losing
2	2	0	0	0
3	6	0	0	0
4	13	1	1	7
5	19	0	0	0
6	17	1	1	5
7	30	0	0	0
8	41	0	1	2
9	57	2	2	3
10	52	2	2	4
11	47	4	13	20
12	35	0	13	27
13	42	5	5	10
14	33	3	8	18
15	16	5	14	40
16	16	9	11	30
17	11	6	12	38
18	16	6	18	45

compared, 91 per cent of the lower grade cases that lose occur after the age of ten, while 94 per cent of the higher grade cases that lose occur after the age of ten. It seems, therefore, that age is also a factor determining deterioration quite independently of grade.

Constancy of the Intelligence Quotient

The method of determining the continuous course of the I. Q. from seven to twenty years from results that are limited to four year periods of examinations and re-examinations for any given case has already been explained in presenting the mental age growth curves. Table IV gives the I. Q.'s that correspond to the mental ages given in Table I. In plotting the average, continuous I. Q. curve from these figures, however, it becomes necessary to take into account the variations between the different groups of cases whose first examinations occurred at given ages. The irregularities in the average curves are thereby at once explained. Thus, in the moron grade especially, the cases examined first at seven happen to be improving cases, and their I. Q.'s cause a rise in the average I. Q. curve from seven to nine.

These results determine certain tendencies of the I. Q. It decreases with increasing age for feeble-minded. That decrease is larger for the high grade than for the low grade cases. Nearly every group examined first at a given age shows both these facts. The averages, given in the last figure under each age, showing the continuous course of the I. Q. to age twenty, reveal them more strikingly. That the I. Q. of the feeble-minded would decrease with age was anticipated, but that it would decrease more for the higher grade than for the lower grade cases was contrary to expectation. I shall show later that both these tendencies of the I. Q. follow mathematically from the fact of a certain type of decreasing rate of growth with increasing age of average normal children, when that growth is measured in absolute units of measurement instead of by the variable unit given in mental ages.

The above figures do not decide the question as to whether the decrease in I. Q. takes place at a uniform rate for all ages for a given grade, or whether it decreases more rapidly at some ages than at others. The rise in the I. Q. beyond the age of fifteen is due, of course, to the fact already noted that the mental age continues to increase beyond this age, while fifteen was the highest age by which the mental age was divided to get the I. Q. Not considering ages beyond fifteen it seems from these figures that there are no very marked changes in the rate of decrease of the I. Q. for different ages, and that the course of the I. Q. is roughly a straight line. A different method of treating the results, given below, however, shows some unquestionable age effect. Taking that part of the average I. Q. curve for each grade that is not obviously disturbed by accidentally varying groups, gives the following average yearly decline in the I. Q. for each of the four grades.

Borderline	2.19 points[9]
Morons	1.21 "
Imbeciles	1.04 "
Idiots	.37 "

[9] In studying the results to determine this average yearly decrease in the I. Q. several different methods were used for eliminating the disturbing factors discussed at the beginning of this article. These methods gave somewhat different figures, but all showed that the I. Q. decreased more with the higher grades than with the lower grades. Even when no attempt was made to eliminate these factors, and when the same range of I. Q.'s was attributed to a grade at all ages, and the average I. Q. computed at each age as found in the raw results these figures were, for borderline 2.2 points yearly loss, for morons 1.3, for imbeciles .5, and a negligible change for idiots.

TABLE IV

IDIOTS

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6		20.00	20.84	21.61	22.17	22.67								
4			20.84	21.70	22.56	21.06	19.56							
5				21.69	21.49	21.29	20.85	20.40						
11					22.07	22.12	22.16	21.71	21.25					
2						21.79	21.37	20.95	19.83	18.71				
3							20.99	20.76	20.52	20.27	20.02			
2								20.96	20.81	20.66	20.56	20.46		
3									20.60	20.27	19.93	19.13	18.33	
										19.98	20.17	19.80	18.33	

IMBECILES

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
8	40.00	39.63	39.26	39.10	38.93									
14		39.63	39.20	38.77	37.63	36.49								
16			39.23	37.73	36.23	35.70	34.17							
21				38.53	35.20	36.63	35.97	35.30						
24					37.00	35.05	33.09	31.28	29.46					
18						35.97	34.14	32.30	31.69	31.08				
25							34.34	33.64	32.94	32.24	31.54			
11								33.13	32.54	31.95	31.54	31.13		
12									31.66	31.64	31.62	31.75	31.87	
14										31.73	31.38	31.02	31.91	32.80
10											31.52	28.87	26.22	25.67
7												30.69	31.19	31.68
													30.30	30.05

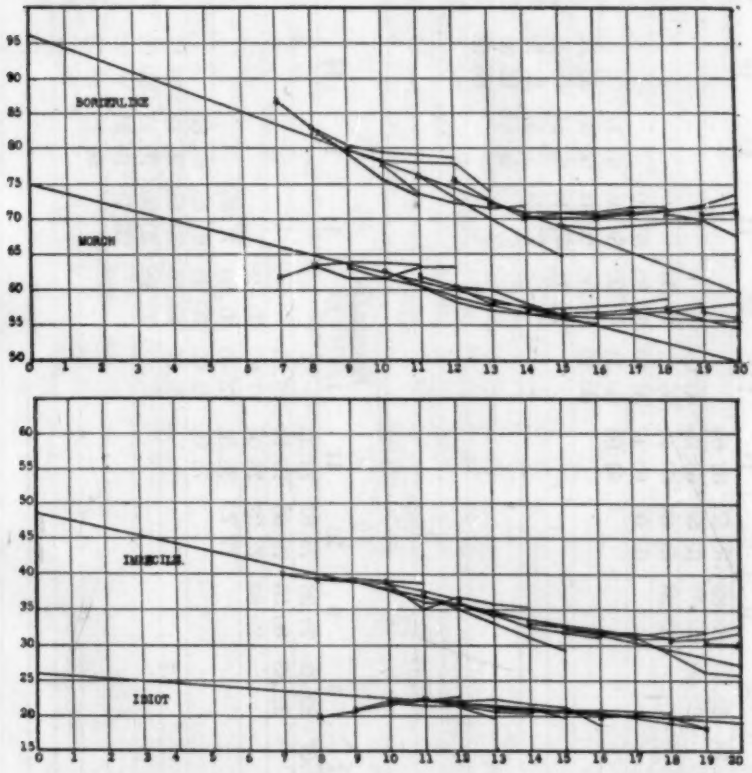
MORONS

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	61.73	62.80	63.88	63.55	63.60									
10		62.80	62.67	62.58	63.25	63.05								
31			63.28	61.68	60.40	59.02	57.59							
24				62.60	60.47	58.36	57.20	56.70						
25					61.93	60.31	60.00	58.08	56.37					
21						60.19	58.06	56.84	56.03	55.21				
17							58.21	57.34	56.51	56.23	56.02			
24								57.24	57.51	57.79	58.20	58.66		
17									56.61	56.98	57.25	57.56	58.20	
17										56.55	56.86	57.17	57.81	58.20
14											57.08	56.09	55.93	55.69
13												57.37	56.09	54.83
													57.01	56.24

BORDERLINE

No.	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6	86.67	82.83	79.33	75.67	73.33									
9		82.83	80.50	79.50	79.16	78.94								
10			79.92	78.02	78.06	77.92	74.22							
10				77.73	73.63	72.40	71.60	71.80						
5					76.05	73.05	70.05	67.45	64.85					
5						75.58	72.78	70.18	70.18	70.18				
12							72.16	70.99	70.91	70.60	71.02			
10								70.11	70.81	70.90	71.71	71.82		
5									69.19	68.79	69.19	69.39	69.39	
4										70.12	70.87	71.02	72.02	73.52
5											70.70	71.30	71.90	72.50
8												70.88	69.75	67.58
													70.76	71.20

The straight lines in the following graphs approximate very closely this average rate of decline, assuming that it is the same for all ages for a given grade. It will be seen in a moment that this assumption cannot be entirely correct. In these graphs the course of the I. Q. for each group of cases examined first at a given age is plotted separately. The small crosses mark the course of the average I. Q. curve. It is seen that this average curve misrepresents the more probably true curve in several instances, most markedly for the beginning of the average curves for the idiot and moron grades.



Relation of Rate of Decline of the Intelligence Quotient to Age. It was seen already that the percentage of cases whose mental age decreases as they grow older increases with age. So far as this might be a general tendency for all cases this effect of age should be more marked for the I. Q., for the

I. Q. may decrease considerably with age without causing an actual loss in mental age. Table V gives the percentage of cases at each age whose I. Q. decreases. The results for all grades are combined, and include no eliminations or corrections for the disturbing factors considered in connection with some of the other questions already discussed. Correcting the mental ages for errors in the scale was found not to make any material difference. The I. Q.'s for intermediate years between two successive examinations of a case were computed and these computed I. Q.'s were used for the results in this table.

TABLE V

Age	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	14–15
Nunber cases	15	42	58	71	122	151	186	233	215	163	182
% decreasing	20	19	45	49	60	71	76	78	75	67	66

These figures show an unquestionable age effect on the decrease in the I. Q. that is not revealed in the figures of the preceding table. The frequency of cases losing in I. Q. increases with age up to about eleven years, the age at which the frequency in mental age loss was found to increase rather suddenly. When the results are computed separately for each of the four grades each grade shows an age at which the frequency of loss in I. Q. is at its maximum and beyond which it declines again, just as for the results of all grades taken together. For idiots this maximum age is twelve, for imbeciles eleven, for morons ten, and for borderline cases seven and twelve. The average mental ages for these grades at these maximum ages are 2.6, 4.1, 6.3, and 6.1 and 9.1, respectively. This relation of frequency of I. Q. decrease to age is therefore not the effect of errors in the mental ages due to the scale of tests. If this were the case the maximum frequency in loss in I. Q. should have occurred approximately at the same mental age. When the mental ages are corrected for these errors, as could be done for mental ages of four and over, the number of cases left for each age becomes rather small for reliability. It makes no material difference for the borderline cases, but raises the maximum age for the morons from ten to thirteen. The correction should, of course, have made more difference for the borderline group than for the moron. We may conclude with safety only that the frequency of loss in I. Q. increases with age, irrespective of grade, up to a certain age. Whether or not this frequency declines beyond a certain maximum age is not so certain.

Relative Frequency of Loss and Gain in Intelligence Quotient. One of the chief questions concerning the I. Q. has been the possibility of predicting future and final mental ages on its basis for the individual case. If the present I. Q. has been determined for a case in a given single examination, how well can his mental age for any age in the future be predicted? Unfortunately this question has been confused with the question of the constancy of the I. Q., with which it has no necessary connection. The ability to make this prediction does not, of course, depend on its constancy but on the regularity of its change in successive examinations at different ages, if there is a change. The traits of the I. Q. established by the present results have complicated its use as a means of prediction. It decreases with age, and the amount of decrease for any given year is dependent on the two further factors of age and grade. Accurate prediction requires that we know not only the presence of these tendencies, but also their extent at each point. But, having discovered these tendencies, we can make corresponding allowances in predicting future mental ages, and thereby reduce a general tendency to error in prediction that would otherwise be present. The following frequency distribution results on the loss and gain in I. Q. gives a fair idea of the general reliability of the I. Q. when used for prediction without allowing for the tendencies to change that was found. Table VI gives the number and percentages of cases that lost or gained 1, 2, 3, 4, etc., points a year. In this table the results of all examinations are again included without eliminations or corrections. They are grouped according to the four grades, and irrespective of age. This, of course, introduces the error into the classification resulting from attributing the same range of I. Q.'s at all ages to a given grade. Since we now know the general tendencies of the I. Q. to change with age and grade the effect this may have on the distribution will be understood. The task of determining the true range of I. Q.'s for each grade at different ages, taking the decline with age of the I. Q. into account, is too complicated for satisfactory solution. Our previous task of determining what cases should be classed under each grade at different ages involved only getting approximately correct average I. Q.'s at each age, which was a relatively simple matter. It is evident that the present procedure cannot give a materially different distribution from what would result if the error in question were eliminated.

TABLE VI

		Over −10	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0
Idiots....	No.								1	6	25	35	48
	%								.6	3.6	15.9	22.6	30.5
Imbeciles.	No.					2	5	6	23	44	178	235	179
	%					.2	.6	.6	2.7	5.4	20.6	27.2	20.6
Morons ..	No.		3	3	11	3	15	49	71	119	133	149	100
	%		.3	.3	1.2	.3	1.5	5.0	7.3	12.2	13.6	15.3	10.2
Borderline	No.	5	10	6		6	43	34	34	23	25	22	11
	%	1.8	3.6	2.2		2.2	15.6	12.3	12.3	8.3	9.0	8.0	4.0

TABLE VI—Continued

		+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	Av.
Idiots.....	No.	19	9	9		3				2		−.06
	%	12.1	5.7	5.7		1.8				1.2		
Imbeciles.....	No.	114	33	23	5	2		2				−.55
	%	13.2	3.8	2.7	.6	.2		.2				
Morons.....	No.	95	75	65	34	13	4	10	10	2		−.72
	%	9.7	7.7	6.7	3.5	1.3	.4	1.0	1.0	.2		
Borderline.....	No.	10	17	13	5	4	3	2				−2.7
	%	3.6	6.2	4.7	1.8	1.5	1.1	.7				

The following summaries from the figures in this table might serve as a rough guide in making predictions, which can then be much improved by taking into account the central tendencies of the I. Q. to decrease with age as found above.[10]

[10] I hope at some time later to discuss more fully the general problem of prediction, in which the present data will be treated specially with reference to this question.

Idiots		Imbeciles	
Per cent.	Points Change	Per cent.	Points Change
31	1	21	1
65	2	61	2
87	3	85	3
96	4	94	4
97	5	97	5
99	6	98	6

Morons		Borderline	
Per cent.	Points Change	Per cent.	Points Change
10	1	4	1
35	2	16	2
57	3	31	3
75	4	44	4
86	5	58	5
93	6	72	6
94	7	88	7
96	8	91	8
98	9	91	9
		94	10

These figures are to be read as follows. Of the idiots 31 per cent change less than 1 point in I. Q. a year; 65 per cent change less than 2 points in I. Q. a year, etc. It is seen from even this rough method of stating the results that much is left to be desired. Yet, we are far from justified in concluding, as is often done, that the I. Q. is of no value at all in making predictions. And further, if predictions are to be attempted at all, what other procedure or method have we that would be even approximately as reliable as predicting on the basis of the I. Q.? Aside from the question of prediction, the I. Q. remains the most accurate and convenient method available for expressing grade of intelligence at any given age of a case.

The Normal Absolute Growth Curve

It remains now to account for the traits of the I. Q. changes with age as found in the present results. The traits to be considered are, (1) that the I. Q. on the whole decreases with age; (2) that this decrease is largest for the highest grade cases studied, those with an I. Q. over .75, and becomes smaller the lower the grade; (3) that loss in I. Q. increases at least in some measure with age.

Considering the general principles of the B.-S. scale, the method of establishing norms and of deriving the mental age and I. Q., there seem to be three possible suppositions that might explain these general facts wholly or in part. The first

is that these traits of the I. Q. result directly from changes in the rate of growth, even when that rate of growth is measured in terms of absolute units. This would mean that the idiot child grows at a retarded rate that kept approximately the same relation to that of the average normal, losing only slightly in relative rate with age. As this retarded rate, however, approaches that of the average normal, giving us the higher grades of subnormals, this loss in relative rate increases, at least up to the grade called borderline in the present study. It is obvious that this loss in relative rate must then decrease again as we approach the average normal rate still closer, since at exactly average normal rate this loss becomes zero and the I. Q. remains constant at 1.00. The correctness of this first supposition is so entirely improbable on the face of it as to really need no further consideration.

A second supposition is that these traits of the I. Q., or at least the first two, may be the result of error in standardizing the scale of tests. It is assumed that the norms for the tests at each age are for non-selected children, representing true averages of all children at each age. This assumption is probably not entirely correct for any scale of tests yet devised. The public schools have always furnished the children in question, and it is recognized that the schools have a selective influence on the average intelligence of the children at each age or school grade. As we pass from younger to older, higher and higher grades of the subnormals are eliminated from the schools through failure to keep up, increasing the average level of intelligence of those that remain. The effect of this process of elimination on the I. Q. with a scale of tests standardized with such so-called non-selected school children would then be as follows: The I. Q. as found would decrease with age for all children of true average intelligence or less, since at each higher age the scale would measure a little bit more too low. It would not increase with age for children above average intelligence. The total decrease in I. Q. for a given period of years, let us say to sixteen, would be the larger the higher the grade of case in question, because the higher grade cases would attain a higher final mental age than the lower grade cases. To make this clear, let us suppose that our scale of tests were standardized in this way for the ages of one to sixteen. Let us assume a group of 100 truly non-selected children at birth, from which one child with a true I. Q. of .70, will drop out each year through some such selection as the schools exert on children in the schools. If norms for tests were established with the

remaining of these 100 children at each age, the I. Q.'s obtained with such tests would be .3 point too low at the age of one, .6 too low at two, .9 too low at three, and so on, to 4.8 points too low at sixteen. With such tests, the average child would lose 4.8 points in I. Q. in sixteen years; the child with a true I. Q. of .75 and attaining a final mental age of twelve at the age of sixteen would lose 3.6 points; the child with a true I. Q. of .50 and with a mental age of eight at sixteen would lose 2.4 points, and so on. Now it is more than probable that the schools exert this selective influence, but there is no data available at present to indicate exactly how much effect this may have had on our norms for our tests. It is possible, therefore, to explain the first two traits of the I. Q. of the feeble-minded, at least in part, on the basis of this resulting error in the scale of tests. I believe that the present results are in some measure due to this factor. The third trait of the I. Q. however, cannot be explained on this ground, nor would it explain an increase in I. Q. with age for children above average intelligence.

The third supposition that may be made to explain the three traits of the I. Q. is that for the average child the yearly increments in mental growth as measured in terms of absolute units decrease each year, giving the general type of growth curve usually assumed to be correct. I shall attempt to show now that this assumption explains not only the three traits of the I. Q. found in the present results, but also the increase in I. Q. found by others for children above average intelligence. This will also suggest certain other traits of the growth curves of subnormals and of the I. Q. that the preceding analysis of the results has not revealed.

In a recent discussion Freeman[11] presents two types of normal growth curves either of which it is claimed would result in the constancy of the I. Q. for cases developing at a subnormal rate. The first is the logarithmic curve, y=log. x. The second is the straight line, assuming a constant rate of development from year to year. For each normal or median curve he plots a mental age curve of a hypothetical case so that at any age the mental age of this case is .66 of the mental age of the median at the same age. That is, the growth curve for the hypothetical case is so plotted in both instances that the I. Q. *will* remain constant at .66. Freeman does not discuss the fact that if the ratio between the median rate of growth and the subnormal rate, measured in absolute

[11] The Interpretation and Application of the Intelligence Quotient. J. Ed. Psychol., Jan., 1921.

units, is assumed to remain constant the logarithmic growth curve cannot give constant I. Q.'s for subnormals. In the case of the straight line type of growth curve for the median, any subnormal developing at a slower than median rate and with a constant I. Q. will also maintain a constant ratio to the median rate of growth, when growth is measured in terms of absolute units, which in this case would be a year of median growth. Now for the logarithmic growth curve this is not true at all. Here these ratios, determined from the y values of the subnormal and median curves at each age, increase with age while the I. Q. remains constant. Thus, for the illustration used by Freeman, the figures run as follows:

Age	3	6	9	15
I. Q.	.66	.66	.66	.66
Ratio	.61	.78	.80	.85

Vice versa, if the rate of growth of a subnormal maintains a constant ratio to the median rate, the I. Q. cannot remain constant, but will decrease with age. Since we have found as a matter of fact that the I. Q. does decrease with age, we must look for the explanation in a median growth curve that will give this decrease in the manner found.

It would be a relatively simple matter to determine the exact nature of the median growth curve if we knew the exact course of the I. Q. from birth to mental maturity, and if we could assume that the rate of development for any grade of intelligence above or below median maintained a constant ratio to the median, when rate is measured in terms of absolute units. We could then construct an absolute median growth curve that would fit the course of the I. Q. changes. We can determine the whole course of the I. Q. from birth to maturity for the grades of cases studied only by assuming that it declines at a uniform rate for all ages. It was seen above that this assumption is probably not very far wrong, the general course of the I. Q. being approximately a straight line, with a small tendency for the rate of decline to increase with age. We may tentatively accept also the assumption that the rate of growth of any grade maintains a constant ratio to the median rate, and then construct an absolute median growth curve that will fit the rate of decline of the I. Q. of one of the four grades of our cases, and with that median growth curve determine the course of the I. Q. for other grades of intelligence. This will show the following. (1) That the absolute median growth rate decreases each year, giving a type of curve similar to the logarithmic, but with the yearly increment in growth decreasing at a much slower rate

than for the curve y=log. x; (2) that this type of growth curve accounts for all the traits of the I. Q. that were established in the above results; (3) that the I. Q. for grades of intelligence above the median will increase with age instead of decrease.

I will take the moron grade, assume that its I. Q. declines at all ages at the uniform rate of 1.21 points a year, which was the average rate of decline found for morons. This gives this grade an average I. Q. of .74 at birth, and of .5585 at the age of fifteen. This gives the following I. Q.'s at the different ages of 0 to 15, with the corresponding mental ages.

TABLE VII

Age	0	1	2	3	4	5	6	7
I. Q....	.74	.7279	.7158	.7037	.6916	.6795	.6674	.6553
M. A...		.73	1.43	2.11	2.77	3.40	4.00	4.59

TABLE VII—Continued

Age	8	9	10	11	12	13	14	15
I. Q..	.6432	.6311	.6190	.6069	.5948	.5827	.5706	.5585
M. A.	5.15	5.68	6.19	6.68	7.14	7.58	7.99	8.38

To construct an absolute median growth curve that will give these I. Q.'s we may set the total growth during the first year from birth to age one arbitrarily at 100 units. Then $1+\frac{\frac{.74y-100}{x}}{2}=.7158$; also, y=100+x, when x equals the increment in growth from one to two years, and y equals the total number of units of growth at age two. Here .74y is the total number of absolute units of growth at age two for this grade of case, and 100 is the total number of units of growth at age one for the median. $1+\frac{.74y-100}{x}$ is then the mental age of this case, which, divided by 2, equals the I. Q. .7158. Solving for x and y, gives 84.31 units increase in

TABLE VIII

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Increase	100	84.31	76.20	67.48	60.65	54.59	48.36	44.66	39.10	36.03	32.02	29.36	26.55	24.75	20.68	
Total	100	184.31	260.51	327.99	388.64	443.23	491.59	536.25	575.35	611.38	643.40	672.76	699.31	724.06	744.75	

growth from 1 to 2 years, and a total of 184.31 units at 2. To get the increase from 2 to 3, we have $2+\frac{\frac{.74y-184.31}{x}}{3}=$ .7037, where y equals the total number of units of growth at 3, x the increment from 2 to 3, and $2+\frac{.74y-184.31}{x}$ equals the mental age at age 3, which divided by the age 3 equals the I. Q. of .7037 at age 3. Here $y=184.31+x$. Solving again for x and y, gives 76.20 units of increase from 2 to 3, and a total of 260.51 units at age 3. Repeating this process, gives the following median growth curve in terms of absolute units of growth, and which will result in the decline in the I. Q. from .74 at birth to .5585 at the age of 15, as found for the moron grade and given in Table VII.

It is seen that this absolute median growth curve decreases from 100 units of growth a year to 20.68 units at the age of 14 to 15. Had either of the other grades instead of the moron grade been chosen and the absolute median growth curve constructed in like manner the general character of this curve would have remained the same. We may next adopt this growth curve tentatively as correct, and determine the resultant I. Q.'s for different grades of intelligence.

Table IX gives the resulting I. Q.'s at age 10 for different grades of intelligence from 1.20 down to .20 in ten point steps. These are derived with the use of the median growth curve given in the figures of Table VIII. It is assumed again that the ratio of total units of development of any grade to the total median units at the same age remains constant. That is, if it is .50 of the median number of units at age 1, it will be .50 of the median at any other age.

TABLE IX

Grade	1.20	1.10	1.00	.90	.80	.70
I. Q. at age 10	1.45	1.20	1.00	.84	.70	.57
No. points change	+.35	+.10	0	—.06	—.10	—.13

TABLE IX—Continued

Grade	.60	.50	.40	.30	.20
I. Q. at age 10	.46	.37	.28	.20	.13
No points change	—.14	—.13	—.12	—.10	—.07

This shows that the median growth curve that results in the I. Q. changes with age as found with morons will give the I. Q. of other grades, in general, as found by our actual

re-examinations of cases. The I. Q. decreases for the lower grades, and more for the higher grades, up to a certain point beyond which this decrease becomes smaller again, reaching zero, of course, for the median grade of 1.00. It also shows that the I. Q. for grades above the median increases with age instead of decreases. The maximum decrease in Table IX is at grade .60, while as actually found in the re-examinations the I. Q. for borderline cases decreased most. The present data are not adequate for determining at just what grade this maximum decrease in I. Q. occurs with the true median growth curve. The median growth curve constructed to fit the moron grade may be more or less incorrect, because (a) the average yearly decline of 1.21 points for morons may be somewhat incorrect; (b) the decline may vary with age more than is assumed here; (c) the ratio of the subnormal rate of growth to the median rate of growth may not remain constant. It can be shown as a matter of fact that this growth curve is probably considerably wrong, by computing the I. Q.'s at each age from 1 to 15 for the idiot, imbecile, and borderline grades, for this median growth curve, and then compare these I. Q.'s with the I. Q.'s of these grades as actually found.

When these computations are made it is found that the present median growth curve gives I. Q.'s that are increasingly higher, from 1 to 15 years, than the I. Q.'s as found for the borderline grade, and increasingly lower than the I. Q.'s as found for the imbecile grade. For the borderline grade, this difference is +.28 points at the age of fifteen. In like manner, the median growth curve that fits the moron grade as found does not correspond with the median growth curve that fits the I. Q.'s of the borderline grade.

This lack of correspondence must be due to one, two or all three of the factors just noted. If the first two factors are not sufficient to give the amount of discrepancy found here, we are forced to the important conclusion that the ratio between median rate of growth and any subnormal rate does not remain constant. In that case it would become necessary to determine the growth curve for every grade empirically by re-examination at successive ages in order to establish the general tendency for each grade in more detail than the present results do.

The median growth curve constructed so as to give the course of I. Q. changes as found for morons cannot show what changes with increasing age in the rate of decline in the I. Q. may take place for other grades, as it is based on a rate of I. Q. change that remains constant at 1.21 points loss

per year. But, having once established the general type of median growth curve that is required to explain the main facts, we may try out modifications of the curve which are constructed in other ways. The present results indicate fairly definitely at what age mental development ceases. The median growth curve used above shows that the number of units of yearly growth decreases at a somewhat increasing rate with age, but it results in no radical modification of the curve to let it decrease at a constant rate. Assuming now a constant rate of decrease in the number of yearly units of growth, and placing the cessation of mental development at the age of sixteen years, gives a yearly decrease in rate of 6.25 units, when the total number of units of growth from birth to age one is placed again at 100. This results in a total of 844 units at fifteen instead of a total of 744 units for the median growth curve used above. Computing the I. Q.'s at different ages for different grades of cases that result from this median growth curve gives a marked increase with age in the rate of decline of the I. Q. for the higher grades of subnormals. Below grade .40 the rate of decline of the I. Q. becomes practically constant for all ages. This is seen in the following figures, omitting fractions.

TABLE X

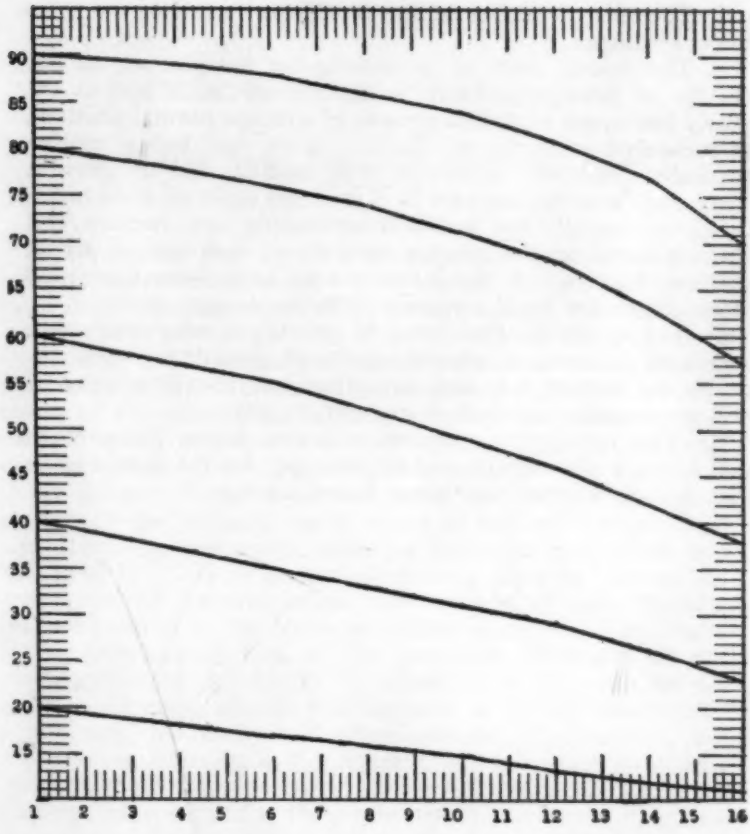
Age		1	4	6	8	10	12	14	16	Av. Yr. Loss
Grade.....	.90	.90	.89	.88	.86	.84	.81	.77	.70	1.31
"	.80	.80	.78	.76	.74	.71	.68	.63	.57	1.53
"	.60	.60	.57	.55	.52	.49	.46	.42	.38	1.47
"	.40	.40	.37	.35	.33	.31	.29	.26	.23	1.13
"	.20	.20	.18	.17	.16	.15	.13	.12	.11	.60

The average yearly loss in I. Q. that results from this median growth curve is given in the last column of figures on the right. It is seen that the maximum decline occurs for grade .80, twenty points higher than for the median growth curve that fits the moron grade I. Q.'s. In other words, a relatively small change in the exact nature of the median growth curve shifts this point of maximum decline in the I. Q. over a wide range. The following graphs are for the figures in Table X.

Summary

1. The mental ages of the feeble-minded increase with age at a rate proportionate to the degree of mental deficiency.

2. On the whole, the mental age ceases to increase between the ages of fifteen and eighteen, the idiot grade ceasing development about three years earlier than the borderline grade.



3. For the cases figuring in the present study and for the time period they were under observation, 4.8 per cent of the cases gain twelve months or more in mental age in a year, 68 per cent gain from one to eleven months a year, 11 per cent gain or lose less than one month a year, and 16 per cent lose from one to seven months a year in mental age.

4. The lower grades lose more frequently in mental age than the higher grades.

5. The frequency in loss in mental age increases with age, independently of grade.

6. The intelligence quotient decreases with age, and more for the higher than for the lower grades. In the present study the average yearly decrease is 2.19 points for the borderline cases, 1.21 for morons, 1.04 for imbeciles, and .37 for idiots.

7. The rate of decline of the intelligence quotient increases some with age.

8. The several traits of the intelligence quotient are all due chiefly or possibly entirely to a decrease with age in the yearly increment in mental growth of average normal children, as measured in terms of absolute units, and follow mathematically from the nature of this median rate of growth. They may in some measure be due to the scale of tests measuring increasingly too low with increasing age, because this scale is based on the average abilities at each age of school children from which the lower grades have been more and more eliminated by the schools with increasing age.

9. It is probable that even if mental development were measured in terms of absolute units of growth the ratio between the median rate and any other rate above or below it will not remain constant from year to year.

10. The intelligence quotient of cases above the average will increase with age instead of decrease, for the same reason that it will decrease for cases below average.

PICTORIAL COMPLETION TEST II[1]

By William Healy, Judge Baker Foundation, Boston

Introduction

It seems clear that any fair attempt to measure intelligence must include well-directed estimations of apperceptive abilities. Surely the capacity for "putting two-and-two together" in the realm of thought, the ability to "turn things over in the mind," to "rationalize" about perceptual material with what may be drawn on from the ideational and other memory stores of the mind is of the very essence of intelligence and of very great importance in civilization. This, if anything, may be spoken of as a higher mental power.

My interest in a pictorial completion test dates from appreciation of the probable great value of testing apperceptions and from the realization that the language completion test (as used by Ebbinghaus and developed since by others) involves several features which interfere greatly with its early intended use as a test for apperceptive abilities. First, there is the outstanding fact of the immense differences in the opportunities of individuals for acquiring facility with the use of our language, not only differences in formal educational advantages, in family and other cultural opportunities, but also, here in America, in the extent to which English is actually spoken at home or—as unfortunately with some of our immigrant peoples wedded to their own institutions—even in school.

[1] This test was exhibited first at the December, 1917, meeting of the American Psychological Association. Soon afterwards the War Department requested the use of it, and, of course, further publication at that time concerning the test became undesirable, even of the scoring values and norms which were at once worked up for the purposes of the Army psychological examinations when the test became part of them. Then the illustration was copyrighted in its several parts in 1918 in attempt to prevent the unfortunate newspaper publicity which befell my first Completion Test. The Judge Baker Foundation (40 Court St., Boston) has recently undertaken to have the test conveniently mounted and boxed and obtainable for a reasonable price.

(Since this article, including the above foot-note, was written, I have observed that this test, including the small plate of the whole series of pictures, which we have purposely refrained from publishing ourselves, and a copy of the scoring values calculated by us, has been published without any mention of authorship in a book entitled, "Army Mental Tests," by Yoakum and Yerkes; Holt & Company.)

On account of this total situation it is impossible to utilize the language completion test as an apperception test equally fair for all.

And whenever the verbal passage becomes really difficult, as in the harder parts of language completion tests, unusual words and those with special shades of meaning increase the difficulties for interpretation of the test results, throwing the advantage strongly to the individual who has had specially good training in language.

Second, I stated my conviction long since, giving instances bearing on the point, that language ability was to be regarded as a specialized ability, sometimes very poorly correlated with the findings on other intelligence tests. We now have considerable data bearing on this. Just as one may possess to a superior or very inferior degree the power of musical expression, so language may "flow trippingly" from the mind or one may by nature halt with words. In the direction of either extreme there is vitiation of the language completion test as indicative of apperceptive ability in general.

Third, the number of details necessary for a desirably complex situation for testing apperceptions can only be presented to a reader at the expense of taking considerable time for the reading and, through this, there is involved another well-known variable, the memory powers or memory span, in this case the memory span for ideas, as they are represented one after another in a long passage.

It suggested itself, then, that a picture completion test might be much fairer as a test for apperceptions, eliminating to a large degree the factors of training and of special ability or disability in language which, as a mode of expression, could be only incidental to the main purpose of the language apperception test. And a well devised pictorial test certainly would give opportunity for presentation of a wealth and exactness of ideational detail at a glance which would require pages of text adequately to cover; as complex a situation as desirable could be depicted without any strain on the memory or any troubled traveling of the attention back and forth, as over the printed page.

In a picture we could present ideas of a simplicity comparable to those of a simple short sentence, or one could illustrate situations too complex for any ordinary rapid solution. And then, as I demonstrate in Pictorial Test II, by the use of the serial pictures one can carry along ideas as they were carried along in the earlier language completion tests where words were elided from a narrative of some length.

Ideas from which to choose in filling the spaces of incompleted sense are represented by a stock of small pictures, and, as one selects verbal ideas from his mental stock in the language completion test, one chooses pictures to fill in what has been cut out of the main illustration.

Development of the Test

The first pictorial completion test which I devised some years ago[2] soon became widely used and commented on, particularly in the monograph devoted to it by Pintner and Anderson.[3] That it seemed to others to embody a sound idea for a good mental test, and that this first picture showed itself to be crude in several features, both for interpretation and scoring, stimulated me to develop plans for a much better test, of the same sort, one that should involve both easy and decidedly difficult problems for apperception. This appeared to be possible by thinking out a pictorial narrative, series of pictures with elements for interpretation carried over from one picture to another just as elements of significance for understanding a narrative are carried over from a sentence or paragraph to later passages, easy or more difficult.

The fact that this idea for the test has not been carried out by others is very likely due to the difficulties and the amount of effort involved.[4] For the development of a series

[2] Described in *The Psychological Review*, May, 1914.

[3] The Picture Completion Test, by Rudolf Pintner and Margaret M. Anderson, pp. 101; Warwick and York, Baltimore; 1917.

[4] All told, a very considerable task has been involved in the development of this second pictorial test, an effort quite beyond the ordinary capacities of any one or two persons. It has been through the fine coöperation of a group, the members of which have given without stint of their time, that the work has been done. Dr. Bronner has been foremost in advice and effort at every stage of the practical development of the test. Indeed, three of the eleven parts of the picture are entirely her suggestions, embodying better ideas than I had been able to advance. The illustrator, Miss Marguerite Davis, brought to the task unflagging zeal as well as artistic skill. The important mathematical suggestions and computations for scoring values were made by Mr. Herbert Sturgis. The entire staff of the Judge Baker Foundation gave itself over to obtaining results on the test both on many trial pictures and on the test as completed, for the establishment of norms. And groups of various nationalities and ages and grades of educational advantages were tested. Much work was undertaken in the interests of the army psychological service during the early part of 1918 in order to establish values for scoring and norms of performance for various groups on this test. Very valuable service in this was rendered by Margaret Fitz and May Bere. A number of outsiders also aided at different stages, a notable amount of work being done by Miss Katherine Coveney.

PLATE I

Sketch from the colored illustration of the simplest picture in the series, the demonstration picture, with the inserts which particularly relate to this picture. Of course only one is quite correct, but the others can be inserted with some show of reasonableness.

of satisfactory ideas a considerable task is demanded of the psychologist, preceding combined work with the artist. One of the main practical obstacles is found in the necessity for devising the correct proportionate size and an insignificant background for all the elided parts, so that there shall be no fitting in of the pieces according to the form of the piece, or by joining together parts of the picture. So far as the size of the pictured object for placement and the immediate background is concerned, any one of the related stock of idea-objects must seem reasonably placeable in any one of the spaces to be filled. Then, too, the carrying over of a central idea from picture to picture with broad illustrative conceptions, but with an exactness of representation that shall satisfy critical interpretation is no easy matter.

Description

Pictorial Completion Test II consists of an illustration to be divided for convenience into two or three parts when mounted. It is composed of a series of eleven pictures 5x3½ inches each (one, very simple in idea, merely for demonstration purposes), representing in sequence situations or events occurring during a day in the life of the boy who is depicted in each picture. From each picture there is a piece cut out, one inch square; the picture to be completed by selection of a square from the accompanying sixty small illustrations, only one of which represents the object which is needed to exactly fulfill the idea of the given picture. In every instance several of the small pieces depict objects designedly logical in part according to whether various ideas suggested by the given picture are taken into account, but not wholly correct, and hence may partially comply with the requirements for thinking out the situation as pictured. A few of the inserts are purposively suggestive. Of course, most of the sixty small pieces are quite inconsistent with the logical needs of any given one of the eleven pictures.

The ten test pictures present situations planned to be of greatly differing difficulties of solution, and they have worked out so well that children of eight years generally get a credit score, while many adults, even of very superior achievement, find it difficult to achieve a perfect score.

The drawing of the illustrations is satisfactory and the bright coloring stands out well as done by the three-color process. The picture is varnished and should be mounted on three-ply scroll-saw wood. The sixty pieces are to be numbered on the back for scoring purposes and are given standard

placings for presentation, according to numbers to be placed on the floor of the box in which they are kept. The pieces which with some reason might belong in any given picture have a certain amount of grouping in our arrangement so that the search for related "ideas" may not be unduly difficult.

(For those who get the illustration not cut up it should be stated that the numbering of the 60 pieces proceeds from the left top across the page and is repeated for each one of the six lines of 10 pictures. Thus, No. 1 is the scissors, No. 10 the picture and frame, No. 60 the top. Of the larger pictures, the boy getting up in the morning is the demonstration piece; No. 1 is the breakfast scene; No. 2, on the way to school; No. 3, at the school house steps; No. 4, the dressing room; No. 5, the schoolroom; No. 6, boys looking up at the sky; No. 7, the street scene with the policeman; No. 8, returning from the baseball game; No. 9, two boys on the side-walk; No. 10, evening scene in the home.)

Directions

The entire picture, parts 1 to 5 forming the top row and parts 6 to 10 below them, should be placed in a good light before the subject. The sixty small pieces in the standardized arrangement are presented in the most convenient place where the subject may see them well.

The procedure is then very simple, but should be according to rule. For very dull persons repetition and increased emphasis of the following verbal directions may be needed.

"Here is a picture—it begins here" (pointing to demonstration picture) "where the boy is getting dressed. It shows the **same boy**, remember, the **same boy**, doing one thing after another during the same day." (Point along the first row, then along the second, to indicate clearly the sequence in which the pictures come.) "You see in each picture a piece is missing. Here" (pointing to them) "are a lot of small pieces, they fit in any of the spaces. But there are more pieces than you can use. The point is to pick out the piece that you think is needed, that is best, to complete the sense of the picture. For instance, What is gone here?" (pointing to demonstration picture). "Yes, a shoe." If incorrect answer is given, which is very rare, Examiner says, "No, he is dressing and he is stooping for his other shoe." "Now, which is the shoe that he must have?" If correct shoe is selected

Examiner says "Yes. This one" (pointing to low shoe) "wouldn't be right because he must have a high shoe to match the other one."

"Now, that is the way each is to be done. There is always some piece that is the very best one; you can tell which it is by studying the picture. Now go ahead."

The main point is to give no help after the first explanation. Nothing further is said while pieces are being inserted except that examinee, if he asks, is told that he may change pieces. After all empty spaces have been filled examiner asks if examinee is quite satisfied, if not, he is told that he may make further changes. When examinee indicates that he has finished as well as he can, time is recorded.

Only the final placings are scored, but record may be kept, if desired, of the changes. Next to the numbers of the pictures in order may be recorded the number or name of the piece inserted.

Twenty minutes is maximum time allowed.[5]

The total score on the test is the sum of the values of the pieces as finally placed. If all are correct the score is 100.

Scoring Values

The value of minus 5 is to be given to all placings where in the table below no numbers are inserted. These placings represent marked absurdities. The numbers in bold type are for correct placings.

Questions will determine in suspected cases whether a correct placing has been accidental. It happens rarely, as when the correct window in IV is placed, although "any window will do." The score of that placing, obviously, should be no more than the verbal judgment calls for, e.g., the value of an incorrect window in IV.

---

[5] In the army examinations only 10 minutes was allowed on account of necessity for saving time. This was rather unfortunate, although, as a matter of fact, it is decidedly rare that longer than this time is desired by the subject. However, the differences that there may be is shown by the fact that one of the most brilliant men in America took the full 20 minutes and wanted more in which to finish Picture 10, finally leaving a piece in place that gave a total score of 90—the relation of the sewing to the meaning of the picture he had not yet apperceived.

TABLE I

VALUES OF PIECES AS PLACED IN PICTURES I TO X

Pieces	I	II	III	IV	V	VI	VII	VIII	IX	X
1					2					12.5
2	0	0	1	2		0	0	0	0	0
3							1	15.5		
4					0					6
5							0	0	0	
6	0			0						
7	0	0	1	2		0	0	0	0	0
8	6			2						
9		5	0		0					0
10	0			1						
11	1			8						
12		0			0					1
13		5	0		3					1
14							1	6		
15			1			1	0	0	0	
16							0	0	0	
17							1	6		
18			0			9.5		0		
19		2			0					0
20										
21		0			1					0
22					2					6
23	1			18						
24					0					6
25							2	0	0	
26			0			4	0	0	0	
27		5	0		1					0
28			0			4	0	0	0	
29										0
30	2			2						
31	1			8						
32		0			7					0
33							1	7		
34							5.5	0	0	
35			11		0	0				0
36			1			2	0	0	0	
37		0	4		0	0				0
38										0
39							0	0	1	
40										1
41		0			2					0
42	0			2						
43		0			2					0
44			1			1	0	0	0	
45	3			2						
46	0	0	6	0	0	0				0
47		5	0		1					2
48							0	0	5	
49		10	0		1					2
	0						0	0	1	

TABLE I—Continued

Pieces	I	II	III	IV	V	VI	VII	VIII	IX	X
51	0	0	1	2	.....	0	0	0	0	0
52	1	.....	.....	8	.....	.....	.....	.....	.....	.....
53	0	0	1	2	.....	0	0	0	0	0
54	.....	.....	.....	.....	.....	.....	0	0	0	.....
55	.....	.....	6	.....	0	0	.....	.....	.....	0
56	.....	.....	.....	.....	.....	.....	2	0	0	.....
57	.....	.....	.....	.....	.....	.....	0	0	1	.....
58	0	0	1	2	.....	0	0	0	0	0
59	.....	.....	.....	.....	.....	.....	0	0	2	.....
60	.....	0	2	.....	0	0	.....	.....	.....	0

The above scoring values were computed from the results obtained by testing 1,542 individuals consecutively, all of them presumably of normal mentality. We did not include records of the feebleminded or psychotic.[6]

The task of working out proportionate scoring values of correct placements upon the basis of all pieces accurately placed, giving a score of 100, was assigned to Mr. Herbert Sturges. He worked first with a series of 257 records and then with 1,542, including the first 257. The resultant values differed very slightly for the two series, indeed, five of the correct pieces remained practically unchanged, less than .5 each, the greatest alteration being from 12.35 to 11.24 as the value of the correct placing in picture 3.

The method applied to the problem was that of weighting each correct solution inversely to the frequency with which it was correctly solved. The working out of the method is as follows: Opposite each frequency write its reciprocals; find the percentage that each reciprocal is of the sum of the reciprocals.

The rule stated is: Value *a* for problem A is the percentage that the reciprocal of the frequency of correct solution of A is of the sum of the reciprocals of all the frequencies of correct solutions.

[6] Here I wish to acknowledge the value of Pintner and Anderson's criticism concerning the inadequacy of the scoring methods and norms for my first completion test. In turn, it was this criticism that challenged Dr. Bronner and me to attempt to develop a less clumsy system of scoring than the one devised by these authors. The main point they make is thoroughly valid; many tests do need standardizing for age and other groups. But on all except group tests the necessary labor—over a thousand hours were devoted to standardizing this our present test—and the requirement for much coöperation stands in the way.

TABLE II

VALUES FOR A SERIES OF 1542 RECORDS

Picture	Correct Placings	Percentage of total correct placings	Reciprocal of frequency	Score	P. E. of Score
1	1130	14.44	0.000885	6. (5.77)	0.06
2	655	8.37	0.001527	10. (9.96)	0.20
3	580	7.41	0.001724	11. (11.24)	0.25
4	363	4.64	0.002755	18. (17.96)	0.56
5	925	11.82	0.001086	7. (7.08)	0.10
6	683	8.73	0.001464	9.5 (9.55)	0.18
7	1228	15.69	0.000814	5.5 (5.31)	0.05
8	416	5.32	0.002404	15.5 (15.67)	0.46
9	1325	16.93	0.000755	5. (4.92)	0.03
10	520	6.65	0.001923	12.5 (12.54)	0.30
Totals	7825	100.00	0.015337	100.00	

The question of giving values to solutions other than the one correct solution of each picture could not be answered by mathematical considerations based on the frequency of such solutions. Arbitrary determinations embodying as much common sense as possible were invoked. No approach to correctness, it was felt, could be worth more than half as much as the correct solution. So the pieces which were planned to represent ideas similar to the single correct piece but missing the real point were given values of half or a little less than half. Thus the value for correct placing, the red book, in Picture 2 is 10 (the nearest half decimal to 9.96). But if one finally chooses the blue book or the two books or the lunch box or the box of pencils the score is 5. Of course it could be nothing else than the red book if one apperceives the situation, the boy having dropped something out of his strap, and his books, plainly colored, being shown strapped together in the previous picture. The only other possibility is a footprint on the ground, and that receives a score of 2. In the other pictures the commonsense of the situation and the possibilities differ and various arbitrary rulings concerning scores have been made accordingly—always based, however, on the original computations of correct scoring values.

If a piece of indifferent meaning, such as a blank in Picture 2, is left in place, or if, as extremely rarely, the space is left unfilled it has been decided that the score should be 0. But a large share of the possible placings are obvious absurdities and logically should be penalized. It is pretty clear just which these absurdities are. It was decided that minus 5 should be the standard score for these absurdities.

Though the scoring of these latter classes of placings is thus numerically arbitrary, still it must be remembered that the scoring as a whole is based upon mathematical computations of fair values according to ascertained relative differences in performance and that once this was done, the minor values as established are as fair for one person as another. (Of course I realize the suggestion value or the association value of any piece is never the same for one person as for another, but the same might be said to a greater or lesser degree about almost any mental test.)

TABLE III

PERCENTILES FROM THE ABOVE CHART

Ages	7	8	9	10	11	12	13	14	15	16	17-20	20-50
75 percentile	24	41	48	59	63	66	69	72	76	75	76	78
Median	9	27	37	47	54	55	58	62	64	66	65	65
25 percentile	−6	7	23	32	41	45	50	52	54	54	54	54

Special educational advantages. Whatever the findings for those who have had educational advantages above high school it would hardly be safe to allege that any good score would be due to such advantages rather than to native ability. It is more than conceivable that a considerable number of those who go further with their schooling do so because of their already demonstrated ability. Nevertheless it was felt to be much wiser to keep separately the 231 records of this group.

TABLE IV

Scores Obtained	No. of Subjects
91–100	57
81–90	54
71–80	47
61–70	42
51–60	21
41–50	8
31–40	2
	231

75 percentile = score of 90
Median = score of 80
25 percentile = score of 68

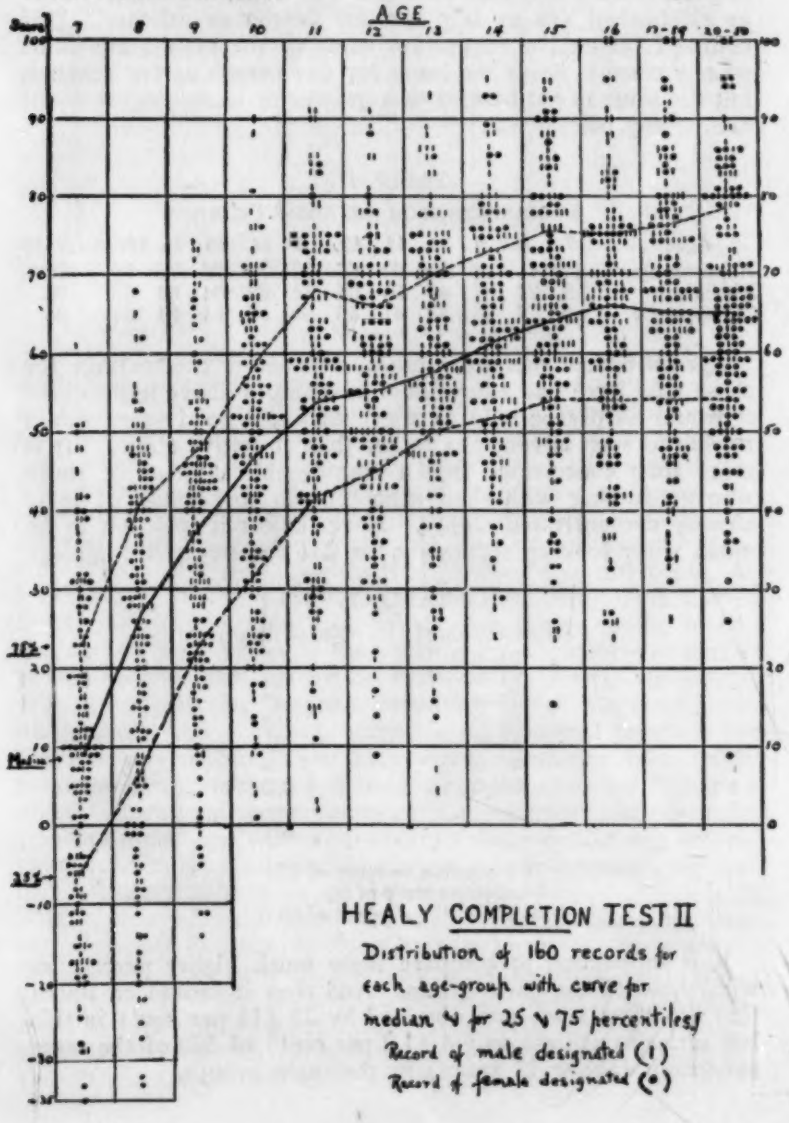
It is interesting to compare these much higher percentiles with those of the main group. And then it should be stated that a perfect score was obtained by 25 (11 per cent) in this last series as against only 4 (1.2 per cent) of 320 of the same age-group (above 17 years) in the main groups.

Norms of Performance

Age-groups

Figure 1

Age distribution of total scores, 160 for each age-group (records of the feebleminded and of those with educational advantages above high school are not included).



While complete significance cannot, of course, be given to the findings for age-periods with the higher educated group on account of uneven and too few numbers, yet it is striking that 20 per cent of 65 individuals 35-50 years of age gave perfect scores and that their median score was 85, while 7.3 per cent of 164 with ages 19-35 did perfectly, with a median score of 76. It would seem as if among well educated people youth did not have the best of it in this test.

The Mentally Defective. Records of the performance of the feebleminded on any one test show the futility of attempting to establish for any single test norms which are characteristic of the mentally defective. Mental defectives differ greatly among themselves in their special capacities and in other important features of mental life than are expressed by general age-level ratings on a given "intelligence scale." Some are feebleminded because of certain defective memory powers, others with good memories have poor reasoning ability, some show great unevenness in their mental powers, some are below par evenly in their various mental capacities. When really good chapters on analytic psychology of mental defectives shall have been written, we may learn something of the complexity that most certainly exists in the mental life of the feebleminded, which to many seems a simple affair. As it stands, we should scan the records of defectives in general on any special test with the knowledge that they represent the performances of individuals with age-level ratings differing widely from actual age experiences and from varying world experiences. Or the records may come from those who are not clearly feebleminded except insofar as social behavior is concerned (institutions for defectives sometimes retain such cases) or from schizophrenics who are not infrequently held among the feebleminded.

Even on a test involving such complex mental functioning as the one under consideration we may find great differences among defectives, at least of the higher grade, who have equal rank on age-level scales. However, the ordinary observation that some show clear mental grasp (apperception) of their own disabilities and various other practical situations, and other defectives of equal age-level rating by the Binet scale or its modifications fail thus to size up the real situation in regard to themselves and other affairs, ought to make the fact of their mental difference in this particular altogether plain and suggest at the outset that their scores on a special test for apperception might differ widely.

Other Groups. From records obtained by way of experiment, but too few for safe deductions, we have noted several suggestions for special studies. Persons who have seen very little of books and are *unaccustomed to observe pictures* understandingly or very little in any way, all of them being, with us, illiterate foreigners, were at a great disadvantage and clearly could not interpret the sky or the interior of a room as such, to say nothing of finer points in the pictures. *Psychotic* individuals have made all sorts of records, as may be supposed—any study of their performance would have to be by types and stages of psychoses. The apperceptive ability of *old age* is a matter of special interest and perhaps might turn out to be of great practical importance, because of the value that is given by tradition and social usage to the opinion of men of age, especially as they function on the bench or in administrative positions.

Discussion

The universal interest evoked by the test is readily understandable. As a series of incomplete pictures it challenges the performance of everyone, for even the mentally defective, above low grades, see enough of the significance of what is depicted so that they become interested and are able to score as having observed correctly and reasoned correctly about some of the material in the pictures. An opportunity like that of everyday life is afforded them for some sort of not altogether unreasonable performance. In this, namely, in developing a test that in given situations for apperception should have simpler elements capable of being readily rationalized by very immature minds and at the same time embodying features of much greater difficulty, up to those relationships which require close thought by competent reasoners, one of the foremost aims of this test has been realized. And it is a fact that even the cleverest, when they note the possibilities of interpretation, enjoy working out the problems of the test as tasks calling for good effort.[7]

The ideas as depicted in the pictures finally drawn are so clear that they present no logically equivocal issues. This has been proved for us not only by the records of performance, but also by the comments of those doing the tests and their acknowledgments as they have later seen the point for correct solution; even after having themselves earlier failed to apper-

[7] We have just heard that the work with French children on this test as done by Marion Pilpel through Simon's clinic in Paris has been very satisfactory. The situations have been readily grasped by these children.

ceive some essentials. Indeed, it is interesting to note that in the two pictures awarded the highest values because of the infrequency of their correct solutions, the correct insert is indicated with more exactitude by objects clearly shown in the main pictures than is the case in any of the other pictures. There is less chance of equivocation in these two pictures because the possibilities of solution, once grasped, are so immediately related to the factual necessities of the situation. The most common failures, then, are directly perceptual and inferential in character, involving faulty observation or faulty reasoning concerning objects and phenomena universally known and among the most plainly observable in the illustrations.

In even the simplest analysis of the test, as a test measuring something much more than primary mental powers, such as memory, we encounter the facts, then, of its easily observable and readily recognizable material, of the facile perceptibility of the elements in and for themselves. The difficulty of the problem comes in rationalizing the value of each whole situation, of the worth of the various elements and of the cut-out objects (the store of ideas) as possible solutions in the given situation—and this is apperception.

Through its presentation of problems that are visual, non-language, ideational, pictorial, a representation of everyday scenes and activities, this test brings into play an important mental ability—whether specialized as being related to observational and pictorial material or capable of wider interpretation—an ability that we would eagerly know more about for many practical purposes. The reaction expressed by one well-known psychologist on first seeing the test, "Well, if I were hiring people to assist me, I think I should like this test to hire them by," may be paralleled by many another opinion of like import from those who have been using this among other tests. Noting that the higher records on the test, we may take the perfect score, 100, for example, have been done by only very occasional youngsters of high ability, as proved by other tests, and by some adults who have established reputations for achievement, and these records have been equalled very rarely indeed by others, for instance, by a domestic cook, we may look further and generally discover in these cases (as in the latter instance where this cook was found to be well known as having exceptionally good common sense and judgment) that there are obtainable evidences of the correlation of such a performance with otherwise known good mental capacities. The powers that this test in some measure gauges are certainly those allied with better human achievement.

THE CONSTRUCTION OF TESTS FOR THE DISCOVERY OF VOCATIONAL FITNESS

By Frank Watts, University of Manchester

It will be agreed that the successful construction of tests for the discovery of vocational fitness must depend in the long run upon an accurate psychological analysis of the various occupations followed by men and women. Only when this has been accomplished in a sufficient number of cases will it be possible to formulate general principles in accordance with which we may know how to proceed in given instances to frame reliable tests for ready use. The attempt to elucidate such general principles may at the moment appear to many to be premature, but it is difficult to spend much time in the work of trying out tests in industry without feeling that the fundamental assumptions underlying the different methods of test-construction call for careful examination. We cannot hope to proceed far in practice with safety without guiding principles; and it is with certain of the theoretical aspects of this subject, therefore, that I propose to deal in this paper. Occasional reviews of the work already done cannot but be helpful to those engaged in actual investigations.

How may we classify the tests already in use? One broad method of classifying them would involve their division into analytic and synthetic tests, the former representing the method of investigating separately the various mental and physical factors which enter into the performance of a piece of work, while the latter type stands for the attempt to measure vocational aptitude as a unitary capacity. Synthetic tests may be further subdivided into the various sample tests such as (1) vocational miniature tests, oral trade tests, picture trade tests, written trade tests and performance trade tests; (2) empiric tests, which at present represent the application of the method of trial-and-error in vocational psychological psychology; and (3) analogous tests which, combining the advantages of the analytic and the sample tests, will, I venture to believe, prove the only satisfactory means in the long run for diagnosing potential as distinct from actual vocational aptitude of a non-routine kind. Let me describe what I conceive to be the advantages and the disadvantages of each

type of test as far as experience permits me safely to do so. I shall venture to take a point of view which is slightly different from that of other workers.[1]

The analytic test, as has been indicated already, attempts to isolate and measure separately in respect of individual applicants for employment those physical and mental factors which are predominantly demanded in the given occupation. There are obvious limitations to the utility of this method. One could hardly employ it to take an extreme example from outside our subject, in determining whether a man (or woman) were suitable for the life-occupation of husband (or wife). Firstly, it would be well-nigh impossible to analyse out into clearly distinct factors all the various mental capacities called for in the state of matrimony. Secondly, if such an analysis were achieved it would be practically useless to attempt to isolate each factor involved for separate investigation and measurement; for how could one be sure of dealing with any one quality with the certainty that it was completely detached from others intimately associated with it? Thirdly, even if these two insuperable difficulties were removed we should still need to know in what proportions the single elements entered into the composition of the complete capacity, and this raises an important problem which I have never seen faced by anyone who has recommended or employed the analytic method of testing vocational fitness. This problem is one of deciding how to "weight" the scores for the various single capacities so that the total score is really a representative score. I have already indicated in another paper that by "weighting" the scores for important capacities in a number of subjects, rankings may be obtained quite different from those obtained when weighting is neglected.[2] Fourthly, to continue our objections, the analytic method assumes that a psychic whole is merely the sum of the factors into which it can be conveniently split up, and this assumption ought not to be accepted in the absence of proof. The belief which would seem to be justified is that personality is not thus arbitrarily to be divided up into part capacities without loss. To

[1] Cf. Essay by Cyril Burt in *Lectures on Industrial Administration*. pp. 79-104. H. Hollingworth, *Vocational Psychology*, ch. v. B. Muscio, *Lectures in Industrial Psychology*, pp. 119-152.

[2] See *The Outlook for Vocational Psychology, Brit. Jour. Psych.*, XI, 2. A simple method would be to allot more marks to some of the tests in any series than to others. As a possible aid to an elucidation of the problem which would then arise of the relative importance of the various capacities entering into a performance a hierarchy of our human capacities is appended.

use a homely illustration, an unanalysaltic factor of experience may result in two women producing two entirely different cakes with similar ingredients mixed in the same proportions, the baking conditions being constant.

Professor Seashore's patient analysis of the requirements of the singer into something like forty different qualities, though an ingenious piece of work from one point of view, would hardly seem, therefore, to be useful to the vocational psychologist in search of a test of singing ability. Separate investigations by scientific methods regarding the range of voice, the ability to sustain tones, the memory span, the delicacy of pitch discrimination, the association type, "mental plasticity," creative imagination, and so on, would prove to be unnecessarily tedious and probably highly unsatisfactory in view of the difficulties already indicated.

Let us take two potential singers A and B who may have happened each to score a hundred marks in a series of tests applied in accordance with the analytic method. Unless they have made identical scores in the separate tests—which is unlikely—does it help us to say they have been "scientifically" proved to possess equal ability or promise? A has scored heavily in one direction because he possesses a magnificent natural voice, but B who lacks such an asset has made up for this deficiency by the possession of marked creative imagination and control of feeling. So that all we can gather from the elaborate tables of results would be, what we probably knew already, that A excelled in one way and B in another.

There seems to be justification, then, for suspecting that in all but the simplest cases the analytic method will probably prove worthless. It is where routine work is concerned and the chief capacities called for are no more than one or two of the sensory or motor functions which are almost mechanical and practically divorced from the higher forms of intelligence, that the analytic test may be useful, and useful here because the intellectual factor does not complicate significantly the various capacities measured. But even then there may be dangers in ignoring in any series of tests the intellectual factor as we shall hope to show later.

The method of "the empiric test" is strongly recommended by Hollingworth (who, however, admits it to be a "rough provisional and unanalysed expedient") as representing "the most promising experimental procedure for the immediate present and perhaps for some time to come."[3]

[3] P. 119, *op. cit.*

It is a method which necessitates no preliminary psychological analysis of tasks. The experimenter simply applies a number of miscellaneous standardised tests, e. g., the *completion test,* the *analogies test,* the *cancellation test,* etc., and notes how closely the scores obtained correspond with estimates of trade ability. Those tests which correlate highly with such trade estimates are then accepted as suitable for diagnosing capacity for the occupation under consideration.

"Perhaps the most perfect example of this purely empirical procedure," writes Hollingworth, "is the investigation which has now been conducted for several years by Mrs. Woolley and her co-workers in Cincinnati. Children who leave the grades to enter directly into some sort of industrial occupation, are examined by a miscellaneous assortment of simple mental tests. These records are preserved and the subsequent successes or failures of the pupils in the various sorts of work undertaken by them in later life are as carefully recorded as possible. It is hoped that when a sufficient amount of material of this nature has been accumulated, the two sets of data may be compared and information thereby secured concerning the relation between ability in the tests and the types and degrees of industrial fitness."[4]

To argue anything of a general nature from this special application of the empiric method is to assume that the interests of boys and girls are either permanent or change in an orderly and predictable manner. But tests which interest the boy or girl while at school and tend therefore to call out a maximum of ability in their performance may cease to attract in later years; so that *if applied later* they must inevitably produce results which are not to be relied upon to the same extent as data for diagnosing vocational aptitude.

But speaking generally, the empiric method is unsatisfactory because it is at best nothing more than a makeshift method for avoiding the difficulties of psychological trade analysis. To continue to employ this method without at the same time making an attempt to psychograph the work concerned is a procedure which will not be commended by any thinking person. The empiric method, more than any other, places us completely at the mercy of *sampling errors* where our human material is concerned, and of the errors, too, of the trade expert upon whom we must rely for trade rankings of the individuals tested. One of the greatest troubles in vocational testing is to find foremen who can give unbiased judgments about the bare vocational aptitude of the workers under them! There is

[4] P. 114, *op. cit.*

always the double tendency present, which few but the most highly trained observers can control, to pass a favorable judgment upon those whose temperaments "fit in" readily with their own so as to make for a harmonious working relationship, and also to see specific aptitude in what Thorndike calls a "halo" of general ability, but never apart from it in its own shape and outline.[5] Thus in the vocational world the empiric method of test construction in effect, shows us the blind leading the blind. Only when we are still completely unsure of our vocational analysis should we be content to employ the empiric method of test construction, and even then not blindly but for the deliberate purpose of opening up a way to a further knowledge of the psychological requirements of an occupation.

The various types of sample tests which have been advocated indicate an attempt to move away from the confusion of theory created by the analysts and the empiricists. But we shall see that if the theorists have unduly neglected the practical requirements of industrial tasks the practical men who have devised trade tests have equally neglected to consider the theoretical implications of their procedure. To take first the commonest form of the sample test which has been used, viz.: the trade performance test. Here we see the applicant for employment given an actual sample of the work which he may later be called upon to perform. His fitness or unfitness is to be gauged from an inspection of the results. This is a method which lends itself readily for use in a large number of occupations; for example, in the choice of typists, clerks, musicians, athletes, skilled artisans, etc. The principal conditions for its successful application would seem to be: (1) that, at least, a representative part of the work is suitable for performance as a test without serious modification and (2) that the results can be expressed quantitatively.[6]

The performance test-method is not without its pitfalls. At first sight it may not seem to call for any great psychological insight and yet to proceed without regard for the influences and effects due to unfamiliarity with the test conditions, lack

[5] In framing intelligence tests Binet solved this difficulty by constructing his *échelle métrique* which provided him with an independent standard for individual measurements. The same principle must be applied to industry, and the writer is at present attempting to do this in the case of one trade.

[6] In Dr. J. C. Chapman's book on "Trade Tests," (Holt & Co., New York, 1921), will be found standardized performance tests for carpenters, pipe fitters, sheet metal workers, electricians, machinists and mechanics, blacksmiths, stenographers and typists, chauffeurs and so on.

of practice, fatigue, the use of unstandardised samples and unstandardised instructions, nervousness, etc., all of which can only be negatived by the employment of scientific methods of procedure is to risk serious error. Moreover, we must know before we can judge a sample piece of work what an average performance is, and this will necessitate considerable research in every case. To be successful, therefore, the sample performance test must be applied under controlled conditions and the results interpreted in relation to known standards. An intimate acquaintance with the methods of experimental psychology will alone ensure that these conditions and standards have been obtained.

The vocational miniature test is a special form of the sample test possessing its general disadvantages with particular ones of its own in addition. It calls upon the subjects of experiment to deal with small "toy" representations of the actual work to be done. Thus a switchboard, much reduced in size, may be used for testing telephonists, or signalmen may be required to manipulate tiny levers in response to demands for signals. It must be at once apparent that the number of tasks which can be reproduced in miniature is extremely small. But the method has been adversely criticised on account of its own intrinsic defects. Münsterberg quite rightly says that "a reduced copy of an external apparatus may arouse ideas, feelings, and volitions, which have little in common with the processes of actual life."[7] There is no defence which can seriously be put up against this criticism.

Moreover, as far as physical labor is concerned there is reason to believe that ability to perform the fine movements called for in working with small models is not at all indicative of ability to perform the larger movements of the actual work which the small model is intended to represent. In the two cases not only will different muscular and nervous co-ordinations be necessary, but also different types of interest. Thus the watchmaker and the miniature-painter would usually be completely unsuited "temperamentally," as we say, as well as physically for employment respectively upon steam-turbines and motor-generators, or upon big poster work. In other words, there is probably no "general factor" in motor ability.

The oral trade test, the picture trade test in which a picture of a tool or piece of work is shown, and the written trade test are all types of sample test based on the questionnaire-method. Such tests obviously, are meant to elicit, and can elicit, nothing more than information. If these tests are used

[7] *Psychology and Industrial Efficiency*, p. 67.

as supplementary to performance tests they may be employed to serve a useful purpose. But most of us harbour the prejudice that to know all about a game of skill or a trade is by no means proof that one is an expert in it, so that if such tests are used as substitutes for performance tests then the general question is raised as to the relation which exists between skill and knowledge, and whether the latter can be accepted as adequately representative of the former. For the elucidation of this problem everything will depend on the type of question asked and the knowledge presupposed in the individual circumstances. In many cases it is, I believe, a mistake to attempt to measure a bare ability independent of experience, because such experience may be the only proof of the presence of the interest upon which permanent efficiency depends. But it would seem on the whole that this type of test is one which can easily be coached for, if we are to take the American Army oral, picture, and written trade tests as truly typical of the method.

The chief objection, however, which the vocational psychologist will be inclined to raise against all sample tests, of no matter what kind, is that they may measure actual rather than potential capacity. They will not help us to select from among individuals without experience of a task those who would benefit most by training. The type of psychological test which is most urgently demanded—the true vocational test, that is, is the test which can be applied to those about to enter an industry with the idea of gauging their probable fitness in advance, so allowing us to advise the misfits to choose another occupation before it is too late. The only satisfactory test which promises well from this point of view (as a means of choosing workers above the routine level) is the skilfully constructed analogous test. The fact that particular analogous tests have not been altogether successful in the past is no argument against their real value.

The analogous test method calls both for a tentative psychological analysis of the work to be done and for the construction of problems calling out as a combination the essential capacities and interests concerned, in much the same proportion as they are demanded in the actual tasks, but in such a manner as to allow potential capacity, when necessary, to compete on equal terms with capacity already fully developed. Really, therefore, it may often become in practice the sample method purged of its peculiar defects. It will be agreed that the work of constructing such tests is no simple matter, but it will prove, I believe, the only real way of progress.

Münsterberg's famous test for the discovery of car-driving aptitude was, up to a certain point, a test fulfilling these requirements. It was defective, not because of any inherent unsoundness in the analogous method, but because it issued from an incomplete analysis of car-driving aptitude. The method was employed by Münsterberg to replace the unsatisfactory method of vocational miniature which had proved defective. "Not the external similarity of the apparatus," wrote Münsterberg, "but exclusively the inner similarity of the mental attitude,"[8] should be the chief characteristic of a test which is intended to reproduce the conditions of work in any occupation. This inner mental attitude is usually so difficult to analyse that it is not surprising that few excellent vocational tests have so far been constructed. Vocational psychography, that is to say, is still litle more than a mere phrase.

What has been attempted so far, however, will enable us, at least, to study our present position. One may say, to sum up the possibilities of the various types of test, that the various sample tests are useful for choosing skilled workers but useless for selecting youths for apprenticeship, that empiric tests should only be used when other methods fail, that analytic tests may be used (subject to reservations) where routine work is concerned, but that where intelligence as well as specific aptitude is necessary the analogous test is needed.

Most of the mistakes and shortcomings characteristic of our early attempts at vocational psychography may now be classified under the headings of errors of inclusion, errors of exclusion, and those errors of proportion already spoken of in connection with analytic tests as responsible partly for their defects.

To consider the errors of inclusion first. Occupations may be grouped into two classes according as to whether intelligence is or is not the predominant factor demanded in them. In a large number of routine occupations the possession of more than an average amount of intelligence makes for dissatisfaction, restlessness, and inefficiency. Yet, especially where the empiric method has been employed, intelligence has been a factor which unwittingly has been allowed to creep into the test performances designed to reveal aptitude for various types of routine work. Tests which are virtually intelligence tests have been applied in various occupations where the workers have happened to be of low grade mental type, so that the sampling factor has obscured the significance of the results. But since differences of intelligence, even if small,

[8] P. 68, *op. cit.*

still exist and so are responsible for the test rankings, a close correspondence between test results and trade ability has been found. But this is not to say that universally wherever we have greater intelligence there we shall have also greater permanent aptitude for the routine work under consideration. For let us suppose that the Knox Cube Test, which has been used successfully to discover capacity among girls employed in label pasting, were to come to the notice of a tyrant with wide powers, ruling a modern community, who had no use for teachers of psychology. Would he be right in acting on the assumption that if these persons endowed with marked introspective habits of mind were given the Knox Cube Test to perform then the most successful would make, for more than a few hours, the best label pasters? Obviously not.

Thus with every routine operation, tests will often work well only supposing that (1) they are tests of aptitudes which are largely independent of intellectual ability, and (2) that the sample from whom we are to make our choice contains no individuals of more than average intelligence and initiative. These conditions were given when Mr. S. E. Thompson [9] carried out his reaction-time experiments with 120 girls employed by him as bicycle-ball inspectors so that he was able to depend entirely on speedy reaction-time as the most essential element making for success in their work. Had the work involved a fair measure of intelligence, mere reaction-times would have been much less significant of capacity. As it was he was able to select 35 girls superior to the others in the essential quality, and by a modification of the conditions of work (better wages, shorter work periods, holidays, etc.), was able to produce an output not less than before.

To put the matter in another way. In many occupations it may be discovered that a specific aptitude or native interest plays a greater part in making for success than intelligence. In such a case we must beware of relying entirely on tests which are really tests of intelligence rather than of specific capacity.

The errors of exclusion are usually due to the neglect of a definitely specific interest in a test-appeal. Thus, Münsterberg's test for car-drivers was excellent in conception but intelligence was allowed to play too considerable a part, while the interest in driving real cars in real streets which alone will ensure permanent satisfaction and efficiency in the work did not enter into the test to any extent. Had the test been one in which only those with the interests already referred to, or

[9] See F. W. Taylor: *Scientific Management.*

at least the capacity for developing them in the actual work, could have done well the College students who performed the test and excelled above the best drivers would have found their proper level.[10]

The foregoing considerations must suggest to the reader how imperative it is that we should attempt to analyse out as completely as possible the various mental and physical factors which enter, or should enter, into the performance of the principal industrial tasks. Now the uncritical acceptance of text book classifications of the mental factors which may enter into the composition of vocational capacity will not help us very much.

In reading works on the subject of vocational selection, one cannot fail to be struck by the fact that writers are apt to use many of the psychological terms for our various aptitudes in such a way as to blind one to their frequent complexity. Thus, among supposed mental "elements," we find such things as tact, perseverance, attentiveness, reliability, caution, accuracy, mental *tempo*, memory, industry, neatness, etc., and the impression conveyed is that there are definite unitary mental factors easily discoverable corresponding to these vague general terms. The result of this use of general terms, however, is frequently unfortunate. For example, when the peasant ploughs a clean straight furrow, or the seamstress shows a beautifully worked hem, or the typist displays one's correspondence with marked legibility, or the housewife maintains a home in spick and span condition, we are apt to say and think that they all manifest in their own way a single indivisible quality called neatness. Such a way of thinking often deceives us, for these varied forms of neatness though conveniently summed up in this manner under a single name are by no means the same qualities psychologically. Only a crank would expect to discover the best ploughman or the best seamstress by means of neatness in a written test. Similarly, if it were agreed that a proof-reader needed to be gifted for detecting errors it would be absurd to test him by any kind of exercise except one dealing with printed matter.

So when one comes to investigate the mental make-up of adult workers as it is expressed in any particular occupation, what is found is that the "elementary" qualities popularly ascribed to such workers are by no means as simple and indivisible as is implied. We do not find as a result of actual experiments that the best proof-readers are always characterised by the possession of a power of close attention applicable

[10] P. 80. *op. cit.*

in many directions, that fitters or instrument-makers display a quick recognition of the significance of form (of no matter what kind), that telephonists develop a good all-round memory, or that street-car drivers exhibit a general ability to apprehend the relations, no matter what, existing between several things, no matter where. What we do find instead is that with the vast majority of adolescents and adults each ability is more likely to be quite clearly specific and limited in scope. That is to say, a person may be a bad observer as regards the common incidents of the streets, but a good observer of natural phenomena; a youth may have a bad memory for dates in history but a good memory for football fixtures; a man may have a remarkably effective power of paying close attention to the movements of machine parts but not to checking figures or letters which pass before him; one may be able to apprehend the relations existing between the many factors involved in a game if one is a spectator, but not if one is a player; a woman may exhibit a remarkably quick mental *tempo* in clerical work, but a deplorably slow *tempo* in organising a home; and so on. Thus it will be clear that no ability can be wholly understood or profitably investigated apart from the living interest which energises it.

The difficulty which will suggest itself to many who have thought about the part which any interest should play in a psychological test is that since most interests do not develop except through the assimilation of knowledge and skill, it is impossible to allow them to enter into the performance of a test without thereby placing their possessors at a great advantage over others equally capable of doing good work whose interests are not yet aroused. It may help to remove this difficulty if we distinguish between the capacity or skill which is determined largely by the existence of an interest or interests which any intelligent person may develop irrespective of his age, and that capacity or skill which is largely due to the development of a specific interest at an early age for the absence of which no amount of intelligence in later life will adequately compensate. Thus if one should be selecting men for responsible open-air work in a distant colony it would probably be much more satisfactory to choose those whose interest in the colonies and in open-air life is of that deep-rooted type which appears early, becomes nourished and strengthened by tales of adventure and danger, and remains proof against the passage of the years. A question such as, "When did you first get to know about this kind of work?" will usually reveal the age and strength of the interest concerned. In many cases

HIERARCHY OF THE HUMAN ABILITIES [11]

	Processes of the Mechanical level	Processes of the Organic level	Processes of the Human level (Man as creature)	Processes of the Human level (Man as creator)
SENSE	The sense of reactions (measurable in speed and acuity); Taste, Touch, Hearing, Sight, Temperature, Kinaesthesis.	SENSORY DISCRIMINATION or sensitiveness to differences in quantity and quality of sense impressions: e.g. heavy and light, smooth and rough, dark and bright, etc.	The finer discriminations; colour, rhythm, pitch, etc.	Ideal sense; of 1. *order* 2. *form* 3. *beauty*
+ INTELLIGENCE		Intuition: *Recognition* of the practical significance of the above difference.	Logical understanding. *Imagery*: e.g. visual, auditory etc.	WISDOM
	Habit Memory.	ASSOCIATIVE MEMORY: for practical experiences. Memory as RECOGNITION.	RECOLLECTION: a. Arbitrary Association. b. Controlled Association.	
+ FEELING	PLEASURE–PAIN.	The Primary Emotions	The Concrete Sentiments with the emotions: joy, pity, sorrow. (When not rational = *prejudices* and *complexes*.)	The Ideal or Abstract Sentiments.
= WILL	IMPULSE: (Appetition and aversion) *HABIT* as shown in	The Primary Instincts (=Native Interests)	ACQUIRED INTERESTS: The Sentiments in action.	The fully enlightened WILL of the 1. saint 2. philosopher 3. artist
	ROUTINE SKILL Attention, as naturally fixating or fluctuating.	ORGANIC SKILL Attention, as determined by native interest.	SCIENTIFIC SKILL Attention, as marked by the ability to concentrate or distribute it in various ways.	ARTISTIC SKILL

[11] See also *Brit. Jour. Psych.*, Vol. XI, Pt. 1, p.

where industrial efficiency depends upon interests developed early in life it will be found possible to construct "age scales" in the Binet manner so that the extent of the progress made in the developing capacity and skill may together be estimated. But even where occupations demand the functioning of interests which normally do not develop early it will still be necessary to appeal to them in our tests since intelligence tests alone will not suffice. We must still contrive to construct tests which will allow the same interest or interests to operate in the test as in the actual work, and the consequence will then be that those individuals possessing such an interest or interests but lacking "all-round" intellectual ability, can compete on equal terms with those intelligent persons who are able to make good at most occupations interesting to the average man or woman.

A COMPARISON OF BRAHMAN AND PANCHAMA CHILDREN IN SOUTH INDIA WITH EACH OTHER AND WITH AMERICAN CHILDREN BY MEANS OF THE GODDARD FORM BOARD

By D. S. Herrick

Do racial differences extend to mental characteristics? Are differences in color and other physical features accompanied by differences in general intelligence, or in psycho-motor ability? These are questions of considerable interest that as regards the people of South India have not been approached from a scientific point of view, nor investigated by scientific methods. Anthropometric observations were made some years ago. The material thus obtained was used in the preparation of a work of several volumes, "Castes and Tribes of South India," by Dr. Thurston. Mental differences were not investigated.

During the latter part of 1919 the writer took up the subject. A series of tests were made with the Goddard Form Board in some 20 or more schools in the Madras Presidency. More than 700 children of ages from four to fourteen were examined and the results tabulated. As it seemed to be worth while to compare not only Indian children with American children of corresponding ages, but also different racial groups of Indian children with each other, half of the tests were made on Brahman children and half on Panchama children. The term "Panchama," it should be explained, signifies in South India the very large social group in which are found the very lowest castes. Hindus do not regard them as castes at all in the strict sense of the word, but rather as "outcastes." It was assumed that, so far as there is any distinct racial cleavage among the native peoples of South India, it would be between the Brahmans and the lowest castes, or the outcastes. A fundamental claim of the Brahmans is that their blood is unmixed with that of other races. While it seems probable that in the early centuries of the Aryan migrations there was considerable intermarriage between the invaders and members of the Dravidian race or races that they found occupying the land, it is fairly certain that for many centuries there has

been no general intermingling of blood. Doubtless the Dravidian tribes were of different degrees of civilization. The most backward of them, the jungle dwellers, of whom there are living examples in many parts of India today, are the progenitors of the outcastes of the present day. So far as intermarriage occurred at all, it must have been with the more civilized of the Dravidians.

That there are marked physical differences between most of the Brahmans of South India and the typical outcaste living in the same village is evident from a casual comparison of their color and features. It may also be shown by a comparison of their cranial and nasal indices. Investigations along this line are to be found in the work by Thurston referred to above. Whether there are also marked differences in mentality that can not be accounted for by differences of environment in its broadest aspect is a question of much interest.

The choice of the kind of test to be used naturally fell upon one of the series of performance tests described in "A Scale of Performance Tests," by Pintner and Paterson. The one that seemed to be the most generally useful was the Goddard Form Board. The Binet system was out of the question, since it has not been adapted for use in any of the numerous languages of India. This form of test depends so much on language responses that it would be hopeless to attempt to get reliable results by means of it before the questions had been carefully translated into the language of the subjects whom it was proposed to test, and proved by actual use. This would have been doubly difficult in the case of the children whom the writer tested, as they belonged to two different language groups. The use of performance tests avoids this difficulty. The Goddard Form Board was chosen as the most practicable test with which to begin the series because it has been so favorably spoken of by eminent psychologists in America. It also seemed to the writer to require the exercise of several kinds of mental ability. Not one of the more than 700 boys and girls tested had ever seen a form board, it is safe to assert. Few, if any, of them in all probability had ever handled blocks of wood or other material of different shapes, much less tried to fit them into holes of corresponding shapes. To be confronted with a board full of holes and a lot of blocks, and to be told to put the blocks into the holes as quickly as possible, was a new situation for each of those children. Thus it was well adapted to test their intelligence. At the same time there was nothing unreasonable in the test, so perfectly simple is it.

The method of giving the test was in general that described in the book by Pintner and Paterson already referred to. The board with the blocks in their respective holes was placed on a table or bench before the subject, who was told to watch closely what the examiner did. The latter then took the blocks out of the holes one by one, and laid them in the standard order in three rows by the side of the board at the subject's right hand. The subject was then told to put the blocks back in their right places. A native teacher was always present to make sure that the examiner's words were understood. In case of doubt he would repeat the command. Three trials were given, and the time and errors (if any) were recorded. The time of the fastest performance was regarded as the index of the subjects psycho-motor ability. In practice it was found best not to ask for speed at the first trial, as that tended to confuse him, and sometimes resulted in wild dashes at the board with little effort to avoid errors. A correct performance was the first thing aimed at. Before the second trial, however, the subject was told to put the blocks in as quickly as possible. Before the third he was urged to his utmost effort for greater speed.

The subjects examined were of both sexes. With a very few exceptions all were attending schools of average excellence. All were town-bred. Thus in the two groups respectively the home life and advantages, were practically the same.

The tests, with few exceptions, were given in the afternoon and generally in an unoccupied room or part of a hall in the school buildings, so that both subject and examiner might be as free as possible from distractions of any kind. Avoidance of distracting sounds was in many places impossible, as can be readily understood by any one who has visited even the best of schools in India, where the hot climate requires that doors and windows be kept wide open. As the pupils are used to these conditions, they need not be regarded as invalidating the tests.

Aside from the effort to secure an equal number of Brahman and Panchama children as subjects, no attempt at selection among the children of the schools where the tests took place was made. If this was done surreptitiously by teachers in any of the schools, it would tend to weed out the dullest ones, since the teachers would wish to have their pupils make as creditable a showing as possible. This would of course result in a higher average performance.

No performance time longer than 81 seconds was included in calculating the medians, in order to exclude records that

suggested decided mental backwardness, and to arrive at norms showing the average ability of ordinary school-going children in South India. Not more than one or two records were thus rejected in any one age group. These were only in the younger ages.

Graph showing median Times of Panchamas, Brahmans, & Americans, from age 4 to 14

Goddard Form Board

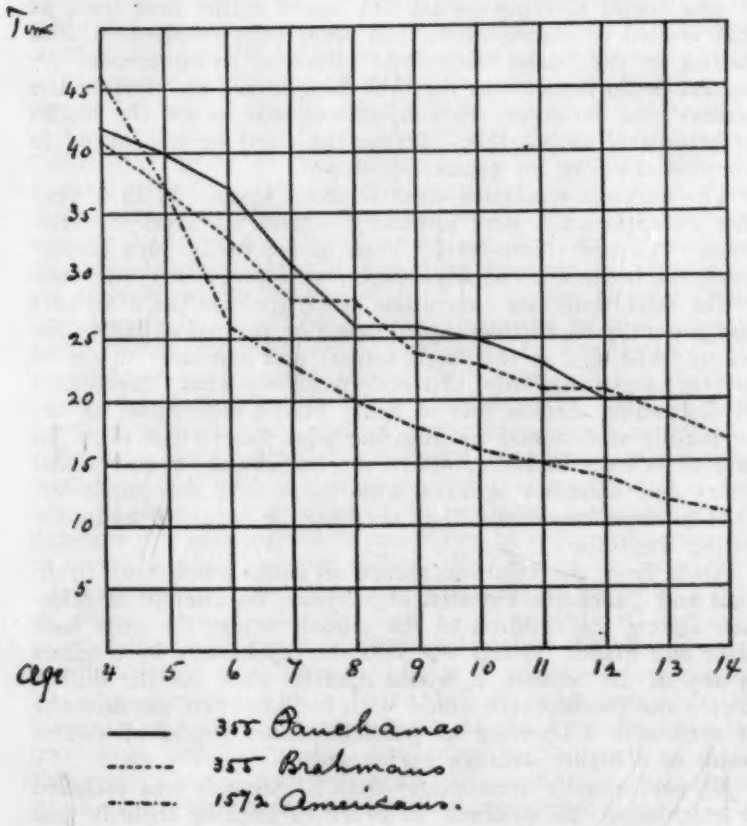


TABLE I—PANCHAMAS

Age	4	5	6	7	8	9	10	11	12	13	14
Time											
14	..	..	..	..	..	..	..	..	..	..	1
15	..	..	..	..	..	..	..	..	3	1	..
16	..	..	..	..	..	..	..	1	2	1	2
17	..	..	..	..	..	1	..	..	3	3	1
18	..	..	..	..	..	2	..	5	5	5	4
19	..	..	..	..	3	1	..	3	5	4	2
20	..	..	..	..	..	..	2	3	4	5	3
21	..	..	..	1	1	1	2	3	5	3	2
22	..	..	..	1	3	2	5	3	3	3	2
23	..	..	..	..	2	2	4	1	3	2	3
24	..	..	..	1	5	3	1	2	3	1	..
25	..	..	..	1	4	3	3	5	1	1	..
26	..	..	1	5	2	3	..	..	..	..	1
27	..	..	1	4	2	4	3	5	1	2	..
28	..	1	1	7	..	..	3	1	..	..	..
29	..	..	1	3	2	1	..	..	1	1	1
30	..	1	2	5	1	1	3	3	..	2	..
31	..	1	1	4	2	3	..	..	..	..	..
32	..	1	1	3	2	..	..	1	..	..	..
33	..	..	2	3	1	..	..	..	..	..	..
34	..	..	2	..	2	1	1	..	..	1	..
35	..	2	..	4	..	3	1	..	..	..	..
36	1	..	4	2	1	..	1	..	..	..	..
37	1	..	5	4	1	1	..	..	..	..	..
38	..	..	1	2	..	..	..	..	..	..	..
39	..	..	1	1	2	..	..	..	..	..	..
40	..	2	2	3	..	..	..	..	..	..	..
41	..	1	1	..	1	1	1	1	..	..	..
42	3	1	..	1	1	..	..	..	..	..	..
43	..	1	1	..	..	..	..	..	..	..	..
44	..	..	1	..	..	..	..	..	..	..	..
45	..	5	1	..	..	..	..	..	..	..	..
46	1	..	..	..	..	..	..	..	..	..	..
47	..	1	..	..	..	..	..	..	..	..	..
48	..	..	2	1	..	..	..	..	..	..	..
49	1	1	..	..	..	..	..	..	..	..	..
50	..	..	..	..	..	..	..	..	..	..	..
51	..	..	1	..	..	..	..	..	..	..	..
52	..	..	..	1	..	..	..	..	..	..	..
53	..	..	..	..	..	..	..	..	..	..	..
54	1	1	..	..	..	..	..	..	..	..	..
55	..	..	1	..	..	..	..	..	..	..	..
56	..	..	..	..	..	..	..	..	..	..	..
57	..	..	1	..	..	..	..	..	..	..	..
58	..	..	..	..	..	..	..	..	..	..	..
59	..	..	..	..	..	..	..	..	..	..	..
60	..	..	..	..	..	..	..	..	..	..	..
61	..	..	1	..	..	..	..	..	..	..	..
62	..	..	..	..	..	..	..	..	..	..	..
63	..	..	..	..	..	..	..	..	..	..	..
64	..	..	..	..	..	..	..	..	..	..	..
65	..	..	..	1	1	..	..	..	..	..	..
66	..	..	..	..	..	..	..	..	..	..	..
67	1	..	..	..	..	..	..	..	..	..	..
68	..	..	..	..	..	..	..	..	..	..	..
69	..	..	..	..	..	..	..	..	..	..	..
70	..	..	..	..	..	..	..	..	..	..	..
71	..	..	..	..	..	..	..	..	..	..	..
72	..	..	..	1	..	..	..	..	..	..	..
73	..	..	..	..	..	..	..	..	..	..	..
79	..	..	1	..	..	..	..	..	..	..	..
81	..	1	..	..	..	..	..	..	..	..	..
Total	9	16	36	59	39	33	30	37	39	35	22
75 Percentile	37	32	33	28	24	23	22	19	18	18	18
Median	42	40	37	31	26	26	25	23	20	20	20
25 Percentile	49	45	43	37	33	30	28	27	22	23	22
Quartile	6	6.5	5	4.5	4.5	3.5	3	4	2	2.5	2

TABLE II—BRAHMANS

Age	4	5	6	7	8	9	10	11	12	13	14
Time	..	..	..	..	..	..	..	..	1	1	..
14	..	..	..	..	..	..	..	1	2	..	6
15	..	..	..	..	..	..	..	..	2	2	2
16	..	..	..	..	..	..	1	1	2	6	4
17	..	..	..	..	1	1	2	8	1	7	6
18	..	..	..	..	..	1	2	5	7	7	2
19	..	..	2	..	1	3	2	2	4	5	..
20	..	..	..	2	2	4	4	2	6	1	1
21	..	..	..	1	3	3	2	2	3	1	..
22	..	..	1	4	1	3	2	4	5	1	..
23	..	..	1	5	2	2	4	1	1	..	1
24	..	..	2	..	3	3	2	3	2	3	..
25	1	..	1	4	3	1	..	5	1	..	..
26	1	3	2	10	3	4	2	1	..	..	..
27	..	1	..	5	4	..	..	1	..	..	..
28	..	..	..	1	1	2	1	1	1	..	..
29	..	..	3	5	5	..	2	..	1	..	..
30	..	1	..	7	..	..	..	..	..	1	..
31	1	..	4	2	1	..	1	..	..	..	..
32	..	1	2	1	2	2	..	..	..	..	..
33	1	..	1	..	..	1	..	..	..	..	..
34	..	..	1	1	1	..	2	..	..	..	..
35	..	1	2	..	1	..	..	..	..	..	..
36	..	1	..	3	1	..	..	..	..	..	..
37	..	..	3	1	..	1	..	..	..	..	..
38	..	1	1	2	..	1	..	..	..	..	..
39	..	1	1	..	1	..	..	..	..	..	..
40	1	..	2	1	..	..	..	..	..	..	..
41	..	1	..	..	..	..	..	..	..	..	..
42	..	..	1	..	1	..	..	..	..	..	..
43	..	..	..	..	..	..	..	..	..	..	..
44	1	1	..	..	1	..	..	..	..	..	..
45	2	2	2	1	1	1	..	..	..	..	..
46	..	1	..	..	..	..	..	..	..	..	..
47	..	..	..	..	..	..	..	..	..	..	..
48	..	..	1	..	..	..	..	..	..	..	..
49	..	..	..	..	..	..	..	..	..	..	..
50	..	1	1	..	..	..	..	..	..	..	..
51	..	..	..	1	..	..	..	..	..	..	..
52	..	..	..	..	..	..	..	..	..	..	..
53	..	..	..	1	..	..	..	..	..	..	..
54	..	..	..	..	..	..	..	..	..	..	..
55	..	..	1	..	..	..	..	..	..	..	..
56	..	..	..	..	..	..	..	..	..	..	..
57	..	..	..	1	..	..	..	..	..	..	..
58	..	..	..	..	..	..	..	..	..	..	..
59	..	..	..	..	..	..	..	..	..	..	..
60	..	..	..	1	..	..	..	..	..	..	..
61	..	..	..	..	..	..	..	..	..	..	..
62	..	..	..	..	..	..	..	..	..	..	..
63	..	..	..	..	..	..	..	..	..	..	..
64	..	..	..	..	..	..	..	..	..	..	..
65	..	..	..	..	..	..	..	..	..	..	..
66	..	..	..	..	..	..	..	..	..	..	..
67	1	..	..	..	..	..	..	..	..	..	..
68	..	..	..	..	..	..	..	..	..	..	..
69	..	..	..	..	..	..	..	..	..	..	..
70	..	..	..	1	..	..	..	..	..	..	..
71	..	..	..	..	..	..	..	..	..	..	..
72	..	..	..	..	..	..	..	..	..	..	..
73	1	..	..	..	..	..	..	..	..	..	..
79	..	..	..	..	..	..	..	..	..	..	..
81	..	..	..	..	..	..	..	..	..	..	..
Total	10	16	36	59	39	33	29	37	39	35	22
75 Percentile	32	27	27	26	24	21	20	18	19	17	15
Median	41	37	33	28	28	24	23	21	21	19	17
25 Percentile	46	45	41	32	32	27	25	25	23	20	18
Quartile	7	9	7	3	4	3	2.5	3.5	2	1.5	1.5

Tables are appended showing in detail the results of the tests.

For convenience of comparison a graph also has been prepared showing the median times of the Panchamas and Brahmans tested by the writer, to which has been added a line showing the median times of American children of corresponding ages according to Sylvester's tests. It will be noted that the Panchamas are from one to four seconds slower than the Brahmans at most ages. The average difference is about two and one-half seconds. Perhaps this can be accounted for by the great difference in the social and educational opportunities enjoyed by the two groups in the past, and by the difference in their environment, which any one acquainted with conditions in India can not fail to have observed.

When we compare the medians of the Brahman children with those of the American group, we find an even more striking difference. At the age of four the American are much slower than the Brahman children, the median times of the two groups being 46 and 41 seconds. At age five, however, the American children have caught up with the Brahmans, the medians for each group being 37. At six years the median time of the American group is 26 seconds, or seven seconds less than that of the Brahmans. In other words, between the ages of four and six the American children have increased the speed of their performance 20 seconds, while the Brahmans have increased theirs only eight. From this point onwards the Americans maintain their lead, and are from five to eight seconds faster at every succeeding age. The average difference is about six and one-third seconds.

The task of assigning causes for this very considerable difference is too large and complicated to attempt fully here. One fact is patent, viz., that the advance made by American children between the ages of four and six is much more rapid than that made by even the quickest group of Indian children of corresponding ages. It is generally accepted that children in the tropics mature earlier than in a colder climate. Presumptive evidence in support of this is found in the fact that the "age of consent," i. e., the age at which a girl may be legally married is fixed by law in India at twelve years. So it is not surprising that the Indian children at age four should be capable of a faster performance than the American four-year-olds.

Again, the gain of 20 seconds in speed of performance made in two years by the American children, as compared with a gain of only eight seconds made in the same period

by the Brahman group, may be accounted for at least in part by the better system and methods of education prevailing in the United States. The kindergarten age lies between four and six. India has not begun to approach America in the thoroughness with which the kindergarten has been developed, nor in the extent to which it has been employed.

Whether these considerations account fully for the differences noted above, or whether the large difference in the median times of the American and the Brahman groups from six to fourteen years of age indicates a fundamental difference in the mentality of the two groups, is a question that need not be discussed here.

One or two other facts of interest are to be noted. First, there is a check at age seven in the case of the Brahmans. They make no increase in speed from age seven to eight, the median times for the two ages being the same. Another occurs at age eleven. There is a similar check in the case of the Panchamas at age eight.

Again, the Panchamas show no increase in the speed of their performance after the age of twelve. So far as the form board shows, they have attained their full mental growth at this age. This, as of course all other inferences based on a single test, would need to be confirmed by other tests.

The quickest time made by any Indian child was 14 seconds, and was made by only three, viz., one Brahman of 12, one of 13, and one Panchama of 14 years. The fastest time made by any American child, as recorded by Sylvester, is nine seconds, and was made by children of 11, 13 and 14 years.

A STATISTICAL METHOD FOR INCOMPLETE ORDER OF MERIT RATINGS

By M. J. Ream, Carnegie Institute of Technology

Rating scales are being used in an increasing number of industrial fields. The diverse situations in which ratings of employees are now being attempted have made necessary certain adaptations of the standard rating procedure. When a number of judges rate a group of individuals by placing them in an order of merit from the best to the poorest, direct comparisons are possible only in case each judge has rated the entire group. But if for some reason, such as lack of acquaintance, some of the judges were able to place in rank order only a portion of the group, the resulting rank positions would not be comparable. A method has been devised for combining such incomplete order of merit ratings in which the judges rate unequal numbers of individuals.

A case in point is the W. T. Grant Company which operates a chain of retail stores throughout the East. This company wished to get estimates of the comparative efficiency of its store managers, but because of their scattered geographical locations, no one judge could know all the men to be rated. Accordingly at a company convention, instructions for rating were given the judges as follows:

"Sheet I attached is a list of all the store managers of the company. First, go through the list and check the names of those with whom you have worked at least six months. Check any others with whom you feel sufficiently well acquainted to judge of their executive abilities. Copy the list of checked names on Sheet II.

"On each remaining sheet, a specific trait is listed and explained. Under this heading write the names you have checked according to merit in this particular trait. Put the name of the best man at the top and the poorest at the bottom....."

The results of this rating contained lists with as many as fifteen men ranked and lists with as few as three ranked. The problem was to combine the estimates of the judges to get the most probable order of merit in each trait.

A survey of the literature on the subject gave but one suggestion, an article by Professor E. L. Thorndike.[1]

[1] Thorndike, E. L. The technique of combining incomplete judgments of the relative positions of n facts made by n judges. *Journal of Philosophy, Psychology and Scientific Methods*, 1916, 13, 197-204.

He suggests that the first step is to get a rough approximation to the true order by mere inspection. The next step is to compare each individual with his neighbor in terms of the percentage of judges (using only those who rated the two individuals in question), who rated the former individual as lower numbered than the latter. Professor Thorndike gives a table in which any given division of opinion among 2, 3, 4, 5, 6,.....15 judges may be expressed in terms of the median deviation. This process of comparison is continued with two or three next neighbors, and the several values determined from the table are averaged. Changes in position are made on a basis of these comparisons.

This method was tried with the Grant Company data but was far from satisfactory. Most laborious was the computation of the *average* differences in terms of the median deviation, from direct comparisons, next neighbor and second and third neighbor comparisons, and from indirect comparisons. After an average was obtained, a change of position was made on the basis of it, and new computations were necessary to justify the new position.

Furthermore, it was possible to get contradictory values by the method of indirect comparisons for individuals who were not directly compared by any judge. The indirect comparisons in addition to being laborious became confusing. Thorndike in his closing paragraph admits the possibility of ambiguity, ".....The resulting order should not without some consideration be used instead of the order derived from comparison with next neighbors only; for the next neighbor comparison, though not utilizing the data fully, is simple and unambiguous in its meaning."

To avoid these difficulties, a new method, or rather a modification of this method was devised and is presented as follows:

(1) For each rank-position on a list containing a given number of names, numerical values are assigned in terms of the standard deviation, assuming a normal distribution in each list of men rated. The values for any given rank-position are shown in Table I. For example, on a list containing six names, position one is assigned +1.49, position two +.68, position three +.22, position four —.22, position five —.68, and position six —1.49. These values were computed from Thorndike's Mental and Social Measurements, Table 22. The value 1.49 assigned to the first position on a list of six names represents the average amount in terms of σ of the upper 16.6% in the group, counted in from the extreme toward the average; the value .68 assigned to position two represents the average

amount in terms of σ of the next highest 16.6% in the group. Values are similarly computed for other rank-positions. When a rank-position is reached which brings the percentage below the average, the σ value to be assigned becomes negative.

It is obvious that an individual ranked at the bottom of a list of three names should be assigned a different value from the individual who is ranked at the bottom of a list of twelve. Table II shows a sample method of tabulating the rating, in this case ten judges on thirteen individuals.

(2) On the basis of these assigned values, an algebraic average is found for each individual, and from these averages a new ranking is made with the individual with the highest average in position one, and the individual with the lowest average in position N.

(3) The verification and correction for this new order of merit is accomplished by the method of simple next neighbor comparisons only. By referring to Table II, the various rankings for any pair of individuals can be quickly obtained. For example, Subjects D and A have been given rank positions 9 and 10 respectively, on the basis of their numerical averages. In this case the two individuals appeared together on the lists of four judges. If Subject D was ranked higher than Subject A in three of the four lists, Subject D would be placed above regardless of the difference in numerical values previously assigned, which in this case was small. Subject D would then be compared with the next neighbor above until he proved to be a loser; similarly Subject A would be pushed down until he proved to be a winner. The simple direct comparison was made the crucial test in any change of position. In cases where the wins and losses were a tie, or in cases where no comparison of the two individuals was made, the numerical averages obtained in the second step above, were the criteria for determining relative positions.

It was found that at the upper and lower extremes of the order of merit, obtained in (2), practically no shifts were necessary as a result of the next-neighbor comparisons in (3), but frequent shifts were needed in the middle of the scale. However, the rank method of rating is generally considered less reliable in the middle of the scale than at the upper and lower ends.

The method just described proved to be much less laborious, and ambiguities and ties could be settled by reference to an independent numerical value. In the case of the Grant Company, the results of the rating differentiated quite strikingly the exceptionally good from the notoriously inefficient store manager.

TABLE I

TABLE OF STANDARD DEVIATIONS VALUES TO BE ASSIGNED FOR RANK POSITIONS

Number of Individuals Rated

	2	3	4	5	6	7	8	9	10	11	12	13	Rank
Rank 1	.80	1.10	1.27	1.40	1.49	1.59	1.67	1.71	1.76	1.81	1.86	1.86	1
Rank 2	—.80	0.	.32	.53	.68	.81	.91	.98	1.04	1.09	1.16	1.21	2
Rank 3	.....	—1.10	— .32	0.	.22	.37	.50	.60	.68	.76	.81	.88	3
Rank 4	.....	.....	—1.27	— .53	— .22	0.	.18	.29	.39	.48	.55	.63	4
Rank 5	.....	.....	.....	—1.40	— .68	— .37	— .18	0.	.13	.24	.32	.40	5
Rank 6	.....	.....	.....	.....	—1.49	— .81	— .50	— .29	— .13	0.	.10	.20	6
Rank 7	.....	.....	.....	.....	.....	—1.59	— .91	— .60	— .39	— .24	— .10	0.	7
Rank 8	.....	.....	.....	.....	.....	.....	—1.67	— .98	— .68	— .48	— .32	— .20	8
Rank 9	.....	.....	.....	.....	.....	.....	.....	—1.71	—1.04	— .76	— .55	— .40	9
Rank 10	.....	.....	.....	.....	.....	.....	.....	.....	—1.76	—1.09	— .81	— .63	10
Rank 11	.....	.....	.....	.....	.....	.....	.....	.....	.....	—1.81	—1.16	— .88	11
Rank 12	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	—1.86	—1.21	12
Rank 13	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	—1.86	13
Rank 14	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	14
Rank 15	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	15
Rank 16	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	16
Rank 17	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	17
Rank 18	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	18
Rank 19	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	19
Rank 20	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	20
Rank 21	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	21
Rank 22	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	22
Rank 23	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	23
Rank 24	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	24
Rank 25	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	25

TABLE I—Continued

TABLE OF STANDARD DEVIATIONS VALUES TO BE ASSIGNED FOR RANK POSITIONS

	Number of Individuals Rated												
	14	15	16	17	18	19	20	21	22	23	24	25	Rank
Rank 1	1.92	1.92	1.99	1.99	2.10	2.10	2.10	2.10	2.16	2.16	2.16	2.16	1
Rank 2	1.26	1.29	1.35	1.35	1.38	1.45	1.45	1.48	1.52	1.52	1.56	1.56	2
Rank 3	.94	.98	1.02	1.04	1.08	1.13	1.15	1.20	1.20	1.23	1.28	1.28	3
Rank 4	.69	.72	.77	.82	.86	.90	.94	.98	1.00	1.04	1.06	1.08	4
Rank 5	.47	.53	.58	.64	.68	.72	.76	.79	.82	.86	.88	.92	5
Rank 6	.27	.35	.41	.47	.51	.55	.60	.63	.67	.71	.74	.77	6
Rank 7	.09	.16	.24	.31	.36	.40	.45	.50	.54	.58	.61	.64	7
Rank 8	— .09	0.	.08	.15	.21	.27	.32	.37	.41	.45	.50	.52	8
Rank 9	— .27	— .16	— .08	0.	.08	.14	.19	.24	.29	.33	.39	.41	9
Rank 10	— .47	— .35	— .24	— .15	— .08	0.	.06	.11	.18	.23	.27	.31	10
Rank 11	— .69	— .53	— .41	— .31	— .21	— .14	— .06	0.	.06	.11	.15	.20	11
Rank 12	— .94	— .72	— .58	— .47	— .36	— .27	— .19	— .11	— .06	0.	.05	.10	12
Rank 13	—1.26	— .89	— .77	— .64	— .51	— .40	— .32	— .24	— .18	— .11	— .05	0.	13
Rank 14	—1.92	—1.29	—1.02	— .82	— .68	— .55	— .45	— .37	— .29	— .23	— .15	— .10	14
Rank 15	.....	—1.92	—1.35	—1.04	— .86	— .72	— .60	— .50	— .41	— .33	— .27	— .20	15
Rank 16	.....	.....	—1.99	—1.35	—1.08	— .90	— .76	— .63	— .54	— .45	— .39	— .31	16
Rank 17	.....	.....	.....	—1.99	—1.38	—1.13	— .94	— .79	— .67	— .58	— .50	— .41	17
Rank 18	.....	.....	.....	.....	—2.10	—1.45	—1.15	— .98	— .82	— .71	— .61	— .52	18
Rank 19	.....	.....	.....	.....	.....	—2.10	—1.45	—1.20	—1.00	— .86	— .74	— .64	19
Rank 20	.....	.....	.....	.....	.....	.....	—2.10	—1.48	—1.20	—1.04	— .88	— .77	20
Rank 21	.....	.....	.....	.....	.....	.....	.....	—2.10	—1.52	—1.23	—1.06	— .92	21
Rank 22	.....	.....	.....	.....	.....	.....	.....	.....	—2.16	—1.52	—1.28	—1.08	22
Rank 23	.....	.....	.....	.....	.....	.....	.....	.....	.....	—2.16	—1.56	—1.28	23
Rank 24	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	—2.16	—1.56	24
Rank 25	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	—2.16	25

TABLE II

RATINGS ON TRAIT X

Subjects Rated

	A	B	C	D	E	F	G	H	I	J	K	L	M
Judge No. 1 (8 comparisons)	—1.67	....	.91	—.91	.18	1.67	.....	—.18	.....	....	—.50	.50	.....
Judge No. 2 (6 comparisons)	.....	.22	1.49	.....	...	.68	—.68	.....	—1.49	.22	.....	....	.....
Judge No. 3 (3 comparisons)	.....	0	.....	.....	...	1.10	.....	.....	.....	....	.....	....	—1.10
Judge No. 4 (10 comparisons)	—.39	1.76	.68	—.13	.13	....	—1.04	.39	—1.76	....	—.68	1.04	.....
Judge No. 5 (5 comparisons)	.....	....	.....	.....	...	1.40	.....	0	—1.40	....	.....	.53	—.53
Judge No. 6 (7 comparisons)	.....	1.59	.....	—.37	...	....	—1.59	.....	.....	.37	—.81	.81	0
Judge No. 7 (8 comparisons)	—.50	.91	—.18	.....	...	1.67	.....	.18	—1.67	....	—.91	.50	.....
Judge No. 8 (4 comparisons)	.....	1.27	.....	.....	...	.32	.....	.....	—1.27	....	.....	....	—.32
Judge No. 9 (7 comparisons)	—.37	1.59	....	0	.37	....	.....	.....	—1.59	.81	—.81	....	.....
Judge No. 10 (9 compa risons)	—.29	1.71	.....	—.60	...	....	—.98	.60	—1.71	.29	.....	.98	0
Total	—3.22	9.05	2.90	—2.01	.68	6.84	—4.29	.99	—10.89	1.25	—3.71	4.36	—1.95
No. times rated	5	8	4	5	3	6	4	5	7	4	5	6	5
Average	—.64	1.13	.72	—.40	.23	1.14	—1.07	.20	—1.55	.31	—.74	.73	—.39
Resulting rank order	10	2	4	9	6	1	12	7	13	5	11	3	8

THE INTELLIGENCE OF CHINESE CHILDREN IN SAN FRANCISCO AND VICINITY

By Kwok Tsuen Yeung, Stanford University[1]

The Problem

The specific purposes of this investigation are:

1. To discover the general intelligence level of Chinese children in the vicinity of San Francisco by means of the Stanford Revision of the Binet tests;
2. To reveal sex differences, if such exist;
3. To compare the intelligence ratings of Chinese and American children of similar social status.

The Method of the Investigation

The Stanford Revision of the Binet scale was given to 109 Chinese children during the spring of 1921. These children ranged in age from five to fourteen years and were all American born. Of the total number, 62 were boys and 47 were girls. In order to secure unselected groups, all of the Chinese children born in San Francisco or vicinity of the ages nine to ten and twelve to thirteen were tested. This first group numbered 49, the second group 33, the remaining 27 being of various ages from five to fourteen. Most of these children came from parents of the laboring and farming classes, chiefly immigrants from the region about the city of Canton. The parents had probably received education in China approximately equal to that secured by the typical village class of Chinese people. The present economic and social environment of the groups from which these Chinese children came probably resembles Chinese conditions much more closely than American life in the vicinity of San Francisco. The Chinese tongue, customs, beliefs, and manner of worship are all retained. They live in their own exclusive settlement, although the children attend the public schools and receive the same instruction as do the white children. The Chinese children master the English language readily.

In using the Binet tests, the vocabulary tests were omitted as unfair to the Chinese children, thus necessitating a small change in the scoring of the tests. Credit for each test passed was given as 2.4 months for the ages at which there are

[1] The writer is greatly indebted to Mr. G. M. Ruch, research assistant in education, Stanford University, for assistance in preparing this manuscript for the press.

six tests in the scale, and 3.4 months for each test at the ages having eight tests. Other than this change, the procedure was exactly that of the author of the tests.

In addition to the tests of intelligence, ratings of school work estimated on the seven-point scale were obtained from the teachers. The teachers, of course, were not given the test scores until after the ratings had been completed. The coefficient of correlation between tests and ratings was obtained by the Pearson product-moment formula. The Pearson coefficient of variability was also used in treating the data for

The Results

TABLE I

THE DISTRIBUTION OF THE INTELLIGENCE QUOTIENTS OF THE 109 CHINESE CHILDREN OF AGES 5 TO 14 YEARS

I Q	Numbers	Percent
66–75	5	4.6
76–85	20	18.3
86–95	26	23.8
96–105	38	34.8
106–115	13	11.9
116–125	3	2.8
126–135	3	2.8
136–145	1	0.9
Total	109	

The median falls at 96.9, or approximately 97, and the empirical mode within the interval 96–105.

TABLE II

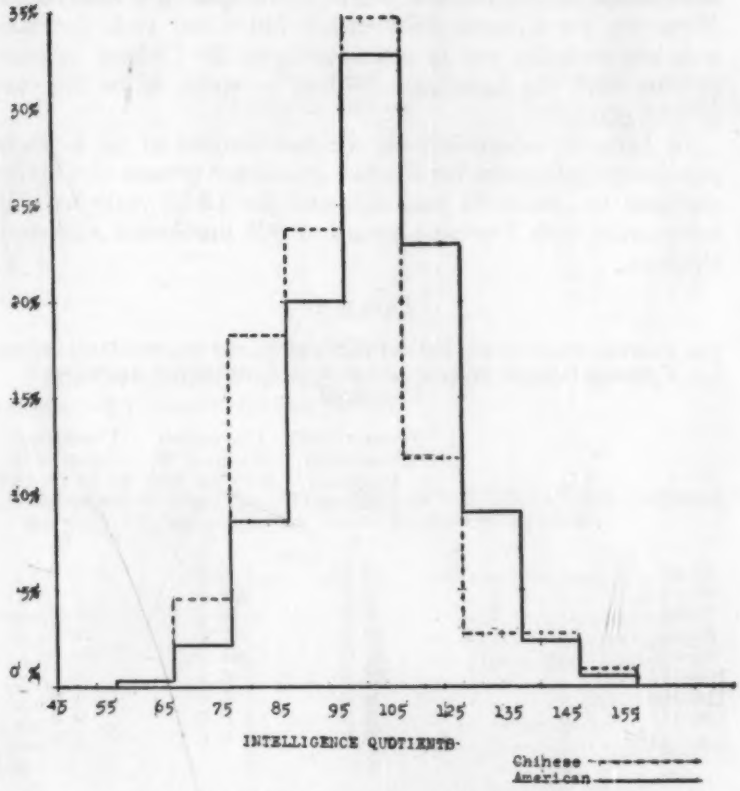
A COMPARISON OF THE I Q'S OF THE CHINESE CHILDREN AND THE 905 UNSELECTED CHILDREN STUDIED BY TERMAN[2]

I Q	Chinese Children	American Children
56–65	0.0	0.33
66–75	4.6	2.3
76–85	18.3	8.6
86–95	23.8	20.1
96–105	34.8	33.9
106–115	11.9	23.1
116–125	2.8	9.0
126–135	2.8	2.3
136–145	0.9	0.55

The median for the unselected American group is 99 in comparison with 97 for the Chinese children studied. This difference as well as the somewhat more restricted range of intelligence quotients of the Chinese group cannot be held to be very significant when the smaller numbers of the Chinese children are taken into account. The possible difference in the amount of selection represented in the two racial groups should also be called to attention in arriving at conclusions from the data presented here.

[2] Terman, L. M.: The Measurement of Intelligence, page 40.

Graph I shows the data of Table II in graphic form.



The following table shows the results of Binet tests given to several racial groups in the vicinity of Stanford University.[8] Except for the Chinese groups, the children were all in kindergartens, who, as stated by Terman, usually show a higher score than do children of higher ages.

TABLE III

Race or Nationality	Numbers	Median I. Q.
Spanish	37	78
Portuguese	23	84
Italian	25	84
Chinese	109	97
North European	14	105
American	49	106

[8] Unpublished data.

It will be seen that the Chinese rank higher than the South Europeans and below the North Europeans and Americans. However, the Chinese children are older and such comparisons are probably not as unfavorable to the Chinese in comparison with the American children as seems to be the case at first glance.

In Table IV which follows, the distributions of the I Q's in percentages are given for the two unselected groups of Chinese children, i.e., the 9-10 year olds and the 12-13 years olds, in comparison with Terman's group of 905 unselected American children.

TABLE IV

THE DISTRIBUTIONS OF THE I Q IN PERCENTAGES FOR THE TWO UNSELECTED CHINESE GROUPS IN COMPARISON WITH UNSELECTED AMERICAN CHILDREN

I Q	Terman's 905 Unselected American Children	Unselected Group of 49 9-10 Yr. Old Chinese Children	Unselected Group of 33 12-13 Yr. Old Chinese Children
	%	%	%
56–65	0.33	0.0	0.0
66–75	2.3	4.1	6.1
76–85	8.6	12.2	21.0
86–95	20.1	26.5	21.0
96–105	33.9	51.0	30.3
106–115	23.1	6.1	6.1
116–125	9.0	0.0	3.0
126–135	2.3	0.0	9.0
136–145	0.55	0.0	3.0

The median I Q for the American, Chinese 9-10 year old, and Chinese 12-13 year old children falls, respectively, at 99.0, 97.4, and 96.0.

The following Table (V) shows the same two Chinese groups in comparison with unselected American groups of the same chronological ages.

The distributions for the Chinese children reveal tendencies for the I Q's to pile up at the average interval 96-105, and to thin out at the upper ranges, in comparison with American children of the same ages. The differences are not strikingly large and the two factors of selection and language familiarity should again be called to attention in interpreting the results.

TABLE V

COMPARISONS OF UNSELECTED CHINESE AND AMERICAN CHILDREN OF AGES 9-10 AND 12-13[4]

I Q	Ages Nine to Ten Years		Ages Twelve to Thirteen Years	
	American	Chinese	American	Chinese
	%	%	%	%
66–75	2.0	4.1	5.0	6.1
76–85	12.0	12.2	15.0	21.0
86–95	21.0	26.0	20.0	21.0
96–105	38.0	51.0	28.0	30.3
106–115	6.0	6.2	19.5	6.1
116–125	1.0	0.0	11.0	3.0
126–135	1.0	0.0	8.0	9.0
136–145	0.0	0.0	0.0	3.0

Sex Differences

Table VI presents the data on sex differences for the total number of Chinese children tested.

TABLE VI

COMPARISONS OF THE SEX DIFFERENCES IN THE 109 CHINESE CHILDREN TESTED, IN WHOLE NUMBERS AND PERCENTAGES

I Q	*Boys*		*Girls*	
	Numbers	Per Cent.	Numbers	Per Cent.
56–65	0	0.0	0	0.0
66–75	4	6.4	1	2.1
76–85	15	24.2	5	10.6
86–95	16	25.8	10	21.3
96–105	16	25.8	21	44.7
106–115	7	11.4	6	12.7
116–125	1	1.6	2	4.2
126–135	3	4.8	1	2.2
136–145	0	0.0	1	2.2
Totals	62		47	

The sex differences are in favor of the girls, the medians being 93.5 for the boys and 99.9 for the girls. The distribution of the I Q's for the girls approaches the normal distribution much more closely than that for the boys, which is skewed downwards, particularily in the 76-85 interval. The extremes of the distributions for the two sexes are not strikingly dissimilar. The ratio between the coefficients of variability (Pearson) for the girls and boys is 0.63. The girls, then, are considerably less variable.

[4] Data on American children from Terman: Stanford Revision and Extension of the Binet-Simon Scale, pages 35, 37.

Teachers' Ratings

In order to make comparisons between the quality of school work and intelligence, ratings were secured from the teachers for the former on the basis of the 7-point scale. These ratings were correlated against the I Q's, giving a Pearson coefficient of 0.52 ± .048. The teachers' estimates correlated with school work (marks) 0.46 ± .060. These values are about the usual ones obtained by many investigators following the same method.

School Progress

A few facts concerning age-grade distributions of the Chinese children are of interest. The 109 pupils were distributed from the first through the eighth grade. Using the standard of ages 6 and 7 being normal for the first grade, ages 7 and 8 for the second grade, etc., it will be noted from the table which follows (Table VII) that the amount of retardation is greater for the Chinese than for the American children, although, with respect to acceleration the comparisons are more favorable to the Chinese. Table VIII gives similar facts when the age-grade classification is made upon the basis of the mental ages. The coefficient between C. A. and grade reached in school is 0.84 ± .028, and for M. A. and grade, 0.88 ± .015.

TABLE VII

ACCELERATION AND RETARDATION FOR THE CHINESE AND AMERICAN GROUPS UPON THE BASIS OF CHRONOLOGICAL AGE IN PERCENTAGES

	Chinese	American
Accelerated		
One year	16.5	14.3
Two years	3.6	1.4
Three years	0.9	0.1
Total	21.0	15.8
Normal	25.0	44.3
Retarded		
One year	38.5	27.9
Two years	12.8	8.8
Three years	1.8	2.4
Four years	0.9	0.6
Total	53.0	39.7

TABLE VIII

ACCELERATION AND RETARDATION FOR THE CHINESE AND AMERICAN GROUPS UPON THE BASIS OF MENTAL AGE IN PERCENTAGES

	Chinese	American
Accelerated		
One year	23.8	15.6
Two years	3.6	3.2
Three years	0.9	0.4
Four years	0.9	0.1
Total	29.2	19.3
Normal	44.9	40.6
Retarded		
One year	19.2	27.2
Two years	5.5	10.2
Three years	0.9	1.9
Four years	....	0.5
Total	25.6	39.6

Inspection of the distribution tables for age and grade will show that by the usual chronological age-grade classification, there is somewhat more acceleration and considerably more retardation among the Chinese. As would logically follow, when mental age is used, there is greater retardation for the American children. In other words, upon a mental age-grade basis the distribution of the Chinese children follows very much more closely the theoretically just basis than with the Americans. The excessive retardation of the Chinese children by the usual standards is very likely due to forced retention of children in grades below those indicated by the mental ages because of unfamiliarity with the English language.

Occupational Status

The occupations of the fathers of the Chinese subjects were obtained and classified upon the 5-fold basis recommended by Taussig. Upon this basis, the great majority of the Chinese children came from parents of classes 3, 4, and 5. Laundry workers and ranchers were most frequent. As has been repeatedly found with American children there is considerable correlation between the I Q and social status, this for the Chinese being $0.49 \pm .049$. The data indicate that the coefficient of correlation is lowered because large numbers of the Chinese children test in intelligence above their social and economic status.

Conclusions

1. No striking differences in the intelligence of Chinese and American children are indicated by the results of this investigation. The Chinese children tested at about the level of Americans and North Europeans and markedly above South Europeans. The median I Q for the Chinese group was 97 in comparison with 99 as found by Terman for the 905 unselected American children studied.

2. Sex differences may be indicated by the data, although the numbers are entirely too small for generalizations. The median I Q for the girls was found to be 99.9 in comparison with 93.5 for boys.

3. Girls were only .63 as variable, according to the Pearson coefficient of variability.

4. Retardation is more frequent among the Chinese pupils than the American, although the Chinese are more perfectly classified in school grade with reference to mental ability than are the American pupils.

5. The Chinese of the vicinity of San Francisco belong chiefly to the lower levels of occupational status although there is a greater observable tendency for mental level to range above economic and social classification.

METHODS FOR THE SELECTION OF COMPTOMETER OPERATORS AND STENOGRAPHERS

By M. A. Bills, Carnegie Institute of Technology
Communicated by C. S. Yoakum

This study was made to determine if certain tests of the Bureau of Personnel Research of Carnegie Institute of Technology would serve first, as a basis for the selection of applicants for courses in stenography and comptometer operation in a technical night school; and second, to determine if from a group of applicants it was possible to select successful comptometer and stenographic operators without an extended tryout. That is, the first purpose demanded that failures be eliminated and the second that sure successes be selected.

Both purposes have been accomplished and the tests are to be used next fall as entrance requirements for the technical school, and as soon as employment is resumed[1] by the firm interested, as guidance in hiring comptometer and stenographic operators.

Both because of the problem presented and because of the methods of attack, which differ from those previously used, the study seems of special interest. The correlation between efficiency in the tests and efficiency in work has been the prime consideration in former studies.[2] In this study, because of the nature of the problem, correlations became merely a by-product, and critical scores the fundamental consideration. Also, a battery of tests has been used rather than a single test, with the idea in mind, that the more measures we could get of a person, the higher would be our ratio of success in selecting promising applicants.

The technical night school where the tests were given was in close connection with a large manufacturing firm and practically all of the teachers held positions in the firm and were therefore well able to judge the quality of work that would be acceptable in the business world. The ages of the students ranged from 16 to over 30. Practically every nationality was

[1] On account of the business depression, the firm is doing practically no employing at the present time.

[2] A Standardized Test for Office Clerks—L. L. Thurstone, *Journal of Applied Psychology*, 1919, 3,—pp. 248-251.

The Value of Mental Tests for Office Clerks—M. J. Ream—To appear soon in Factory.

represented and all classes of home conditions. All students understood English.

Three tests were given to the entire school (139 cases). The tests were:

1. A general intelligence test, designated in this article as Test VI.

2. A test devised to show special aptitude in stenography and typewriting, designated in this article as Test VIII. This test consists of five parts.

3. The Will-Temperament test, designated in this article as Test IX adapted from Downey's test and so revised that it can be used as a group test. This test has ten parts.

The criteria used were the failures and successes in the different classes. The teachers were asked not only to grade the student, but to indicate the line of work for which they considered her best adapted. A person was considered a failure when her teacher did not recommend that she continue work in her course or enter an advanced course. A successful student was one that was recommended to continue her work. For example, if the teacher of stenography recommended that a student take up typing only, that student was considered a failure. If the teacher in typing recommended that a student continue in typing or take up stenography, that student was considered a success, while she was considered a failure if clerical or messenger work was recommended.

Of those tested 84 were in the course in stenography and 35 were in the course in comptometry. The remainder were in a preparatory course and will not be considered. Of the stenographers, 80 of the 84 took the entire series of tests and will be considered throughout the discussion. The other four have been included in the discussion of the tests which they took. Of the comptometer operators, the entire 35 took all three tests. Of the 84 stenographers, 26 were rated as failures. Of the 35 comptometer operators, 14 were rated as failures.

The tests that we found to be most useful in eliminating failures and picking successes were Test VI, Parts 1 and 5 of Test VIII and Parts 2, 3, and 10[3] of Test IX. As a whole it was found that Test VI (general intelligence) and Test IX (will-temperament) were the most successful in eliminating failures and that Test VIII (special ability) was the most efficient in selecting stenographers who were certain to be successful.

[3] These correspond to the parts of the will-profile that Miss Downey has called motor inhibition, speed of decision and perseverance.

Following are the results for stenographers in detail. We first attempted to eliminate failures. Our method can best

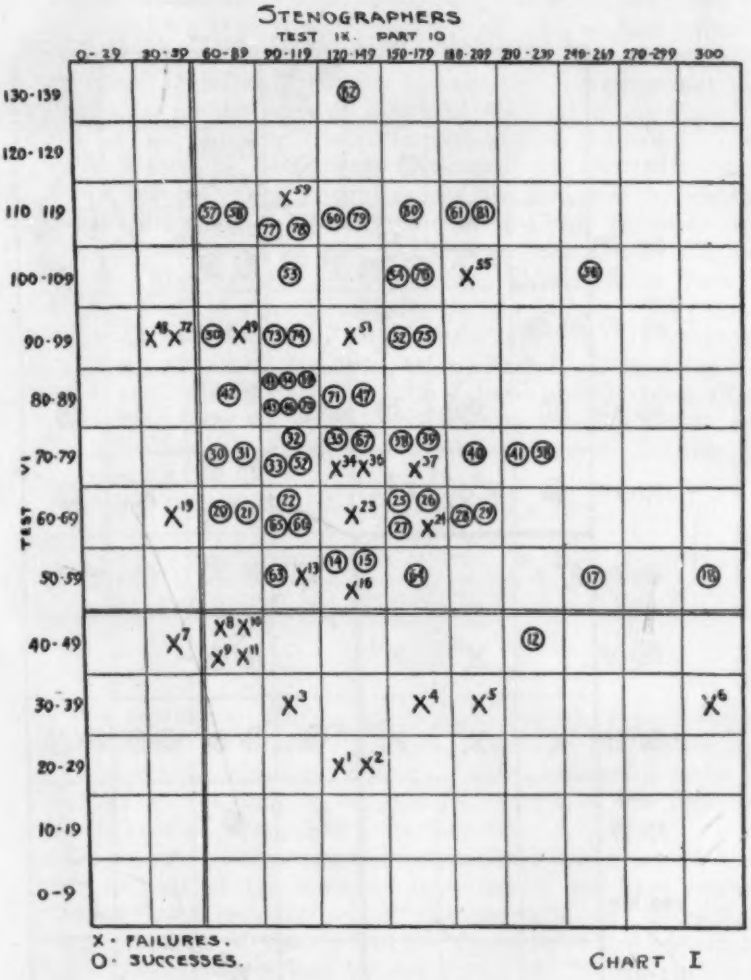


CHART I

be illustrated by charts. In Chart I we have plotted scores made in Test VI against the scores made in Test IX, Part 10. We have indicated in the chart the successes and failures (O standing for successes—X standing for failures). Each X and each O stands for a person; each person has also been

STENOGRAPHERS

TEST IX. - SPEED OF DECISION OBJECTIVE - PART 3

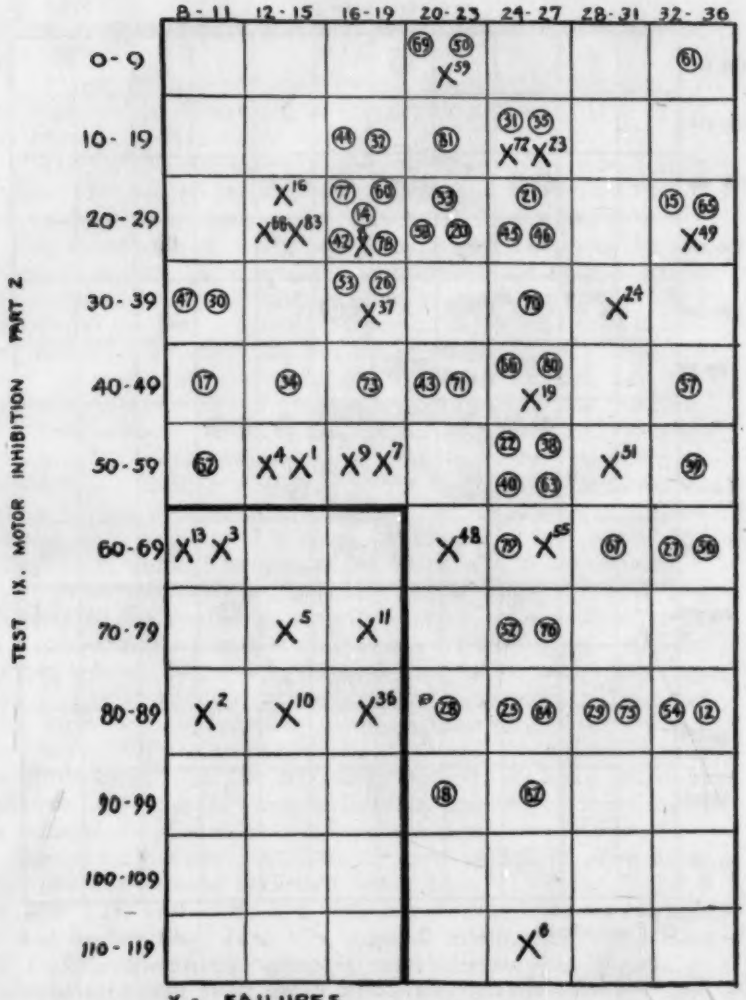


CHART II.

numbered. The numbers below 63 represent Juniors and the numbers 63 and above represent Seniors. It will be seen that 50 seems to be a critical score in Test VI. If we had refused admittance to all scoring below this point we would have eliminated 11 failures and only 1 success.[4] If we continue the elimination on the basis of a critical score of 60 in Part 10, Test IX, we eliminate 3 more failures and no successes. This makes a total of 14 failures eliminated out of the 26, and only 1 success. The critical scores are indicated by double lines on the charts. Now if we turn to Chart II, where we have plotted Part 3 of Test IX against Part 2, Test IX, we find that the combination of a score above 50[5] in Part 2 and below 20 in Part 3 indicates certain failure. This combination would eliminate 7 out of the 26 failures. The persons have been numbered the same as in Chart I. Comparing the two charts, we see that 5 of the 7 have already been eliminated with Test VI, so that, considering the two charts, we find that we can eliminate 16 out of the 26 failures in stenography.

Putting this into a four fold table, we have:

Failed in tests and failed in course.....	16	cases
Failed in tests and succeeded in course.....	10	"
Succeeded in tests and succeeded in course....	56	"
Succeeded in tests and failed in course.....	1	"
Total	83	"

This means that 86% were correctly placed by preliminary examination. It is also important to note in this connection that no person who has become a Senior received a score in Test VI of less than 50, in Test IX, Part 3 of less than 24, and in Test IX, Part 2 of more than 50.

On account of the complication involved in the use of four tests as part of the entrance requirement, we have recommended to the technical school that two tests only be used in the elimination of those from the course who are certain to fail. These tests are Test VI and Part 10 of Test IX. Others are to be given however, for their scientific interest.

[4] Upon personal discussion with the teacher we found that this one success was a doubtful one. The probability that this elimination can have taken place by mere chance is 1 in 26,518 worked out by the ordinary theory of probability, so that we may be reasonably sure of the causal relation between Test VI and success in stenography.

[5] A high score in Part 2 indicates a poor standing.

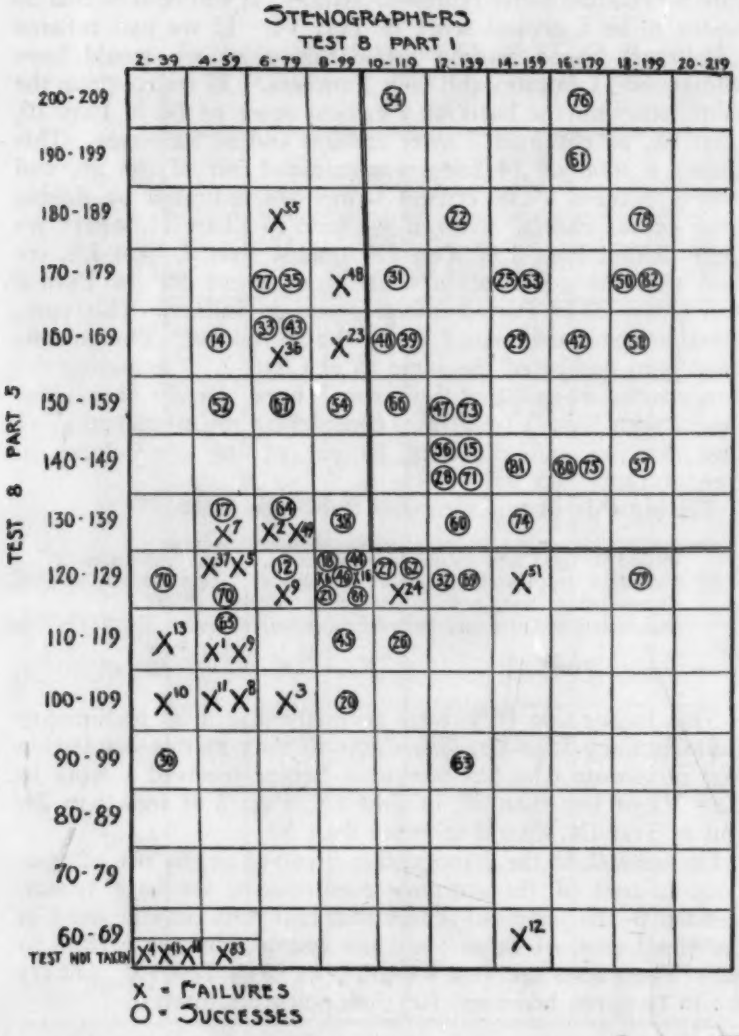

STENOGRAPHERS
TEST 8 PART 1
2-39
4-59
6-79
8-99
10-119
12-139
14-159
16-179
18-199
20-219
200-209
190-199
180-189
170-179
160-169
150-159
140-149
130-139
120-129
110-119
100-109
90-99
80-89
70-79
60-69
TEST 8 PART 5
TEST NOT TAKEN
X = FAILURES
O = SUCCESSES

CHART III.

Second, we attempted to select those from the group who were certain of success. In Chart III we have plotted Part 1 of Test VIII and Part 5 of Test VIII against each other. Again, O's represent successes and X's failures. Each person is numbered as before. The numbers below 63 represent Juniors, and 63 and above represent Seniors. It will be seen that a score above 10 in Part 1 and above 130 in Part 5 indicates certain success. These grades should be used as critical scores in the employing of applicants when a large number are present from whom to make the choice and when no chances are to be taken on the success and failure of the applicant. In every employment office such instances arise. Considering Chart III from a mathematical view point, we find that each part predicts successes and failures as follows:

Part 5—Critical Score, 130

Failed in test and failed in course	15	cases
Failed in test and succeeded in course	19	"
Succeeded in test and succeeded in course	39	"
Succeeded in test and failed in course	7	"
Total	80	"

This shows a correct choice of 67%.

Part 1—Critical Score, 10

Failed in test and failed in course	23	cases
Failed in test and succeeded in course	23	"
Succeeded in test and succeeded in course	35	"
Succeeded in test and failed in course	3	"
Total	84	"

This shows a correct choice of 69%.

A combination of the two parts of Test VIII gives the following results:

Failed in test and failed in course	26	cases
Failed in test and succeeded in course	30	"
Succeeded in test and succeeded in course	28	"
Succeeded in test and failed in course	None	
Total	84	"

This gives 64% correct choice. As will be noted, this correct choice is not as great as for Test VI but the important thing to note here is that success in the tests predicts success in the course perfectly (100%).

If, on the other hand, we take 8 as the critical score in Part 1 and 120 as the critical score in Part 5, we have:

Failed in test and failed in course.....	19	cases
Failed in test and succeeded in course.....	14	"
Succeeded in test and succeeded in course.....	44	"
Succeeded in test and failed in course.....	7	"
Total	84	"

Comptometer Operators

TEST IX. PART 10

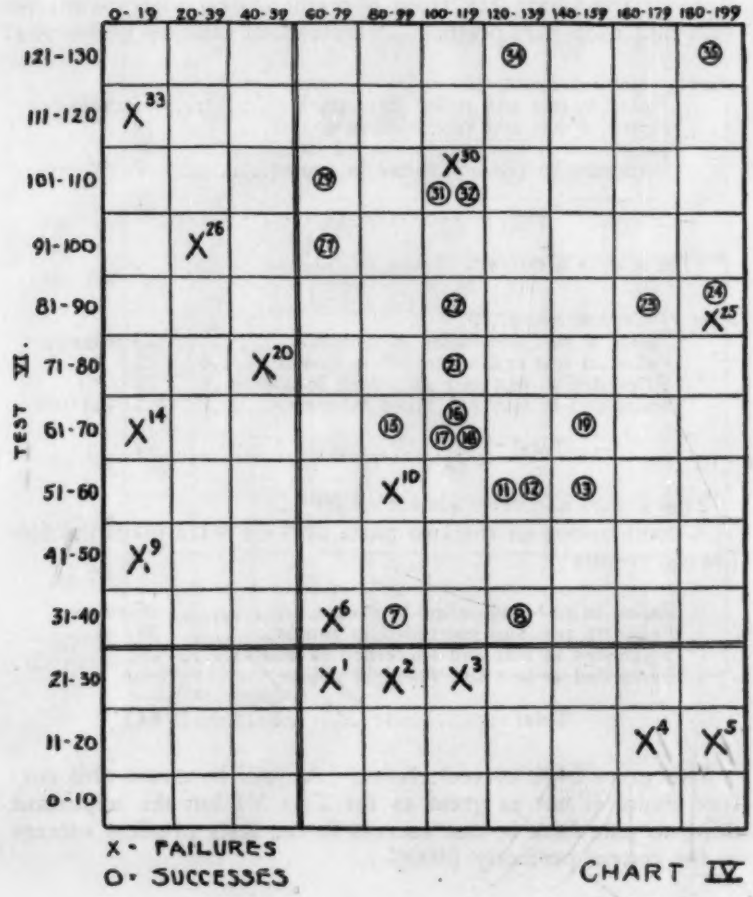


CHART IV

This shows a correct choice of 77%.

In picking our applicants that are sure of success we have, however eliminated some who would have succeeded, and therefore this criterion should only be used when the number of applicants considerably exceeds the openings in the company; or when the selecting of a failure would be a very serious matter.

In Chart IV, we have plotted the scores of the comptometer operators in Test VI and Part 10 of Test IX. It will be seen that a combination of the two tests will eliminate 10 of the 14 failures. The critical scores are 30 for Test VI and 60 for Part 10 of Test IX. The chance that this is a mere chance elimination is 1 in 25,921. We have found no combination of tests for the comptometer operators that would pick successes more favorably than these two tests with which we have eliminated the failures. This is due to the fact that at the present time we have no special ability test for comptometry. We hope in the near future to work out such a test.

Putting the results on comptometer operators in mathematical form, we find that they predict success and failure as follows:

Failed in tests and failed in course.....	10	cases
Failed in tests and succeeded in course.....	0	"
Succeeded in tests and succeeded in course....	20	"
Succeeded in tests and failed in course.....	5	"
Total	35	"

This shows a correct choice of 85%.

Results

I. A battery of tests is more effective both in eliminating failures and picking successes, than any single test.

II. Of the single tests, General Intelligence is the most efficient for eliminating failures.

III. Of single tests, the Special Ability test is most efficient for selecting successes.

IV. Failures can be predicted by the tests with over 85% accuracy.

V. Successful stenographers can be selected.

MINOR STUDIES IN EDUCATIONAL PSYCHOLOGY

By Frances I. Gaw, Boston Psychopathic Hospital, Boston, Mass.

I. School Ratings and Moving Pictures

The educational value of moving pictures and their influence upon behavior have been much discussed questions in the last few years. Educational films of various sorts have become common, and the release of such a picture as the "Fit to Fight Film" in many camps during the war gave official recognition of the educational possibilities of movies. Dr. William Healy in his "Individual Delinquent" in reporting Juvenile Court cases where movies have directly affected conduct, says. "In nearly all people visual memory and visual imagery play the most dynamic part in mental life. . . . When it comes to motion pictures we have added elements of force for the production of either good or bad." The importance of moving pictures as a method of popular instruction is borne out by the existence of such a publication as the Educational Film Magazine.

This work was done to study the influence of movies on the conduct and school ratings of 337 children in a suburb of Boston. In this community there are no moving picture theaters, and the sentiment against them is strong. The children studied were in the 6th, 7th and 8th grades. They were given a questionnaire to determine how many movies they saw a month. Their ratings in studies and deportment were supplied by their school teachers, and correlations were then calculated between these averages and the number of movies attended monthly. The coefficient was found to be, in the case of school studies and movies,—13. The correlation though inverse, is so low that practically no relationship is shown between the two factors considered. In the case of deportment grades and movies attended, the situation is similar with a correlation of —12. The average number of movies attended monthly was .97, or slightly less than one a month. Only 44 children went as often as once a week, and only 6 more than once a week. Over half of the 337 children considered did not go to the movies at all. The two pupils with the lowest averages in their studies were among those who went to movies the most, although many of the other pupils low in their studies and those lowest in deportment went to the movies either an average number of times or not at all. A decided majority of the children with the best averages in both studies and deportment did not go to the movies at all.

The following questionnaire (for which the writer was not responsible) was used in gathering the data for this study:

"1. Do you go to the movies regularly?
2. Do you go to the movies as often as once a week?
3. Do you go to the movies as often as once in two weeks?
4. Do you go to the movies as often as once in a month?
5. Do you go to the movies less frequently than once a month?"

Some children in answering became confused, contradicted themselves and in a few cases added a question. There were 4 or 5 questionnaires so inaccurate that they could not be counted in the correlation. Such a questionnaire might well be modified to read:

1. Do you go to the movies at all?
2. Do you go to the movies as often as once a month?
3. Do you go to the movies as often as once every two weeks?
4. Do you go to the movies as often as once a week?
5. Do you go to the movies more than once a week?

Whether or not this community is typical in regard to the number of movies the children attend, it is obvious that in this study the average movie attendance—once a month—is practically unrelated to either the school work or deportment of the children considered.

II. Relations of Stanford Tests and Dearborn Maze Tests

In this study, the relationship was studied between the Stanford-Binet mental ages of a group of patients at the Boston Psychopathic Hospital and their scores on Test No. 7, General Examination 5 of the Dearborn Group Tests of Intelligence. The Dearborn test used consists of four mazes. It is explained to the subject that the first two are maps of a city which he is to traverse in the shortest way possible, keeping in the streets throughout. The last two are also maps of a city which he is to traverse, stopping at certain points, but he must follow two definite traffic rules: first, he can turn only to the right, and, second, he must always stop on the right-hand side of the street. The third maze is divided into four parts. The first one of these was used as an illustration in the cases examined. Since this study was made, however, the third part also has been used by the author of the test as an illustration. The two more difficult mazes of the test are probably made considerably easier by this change.

In addition to the patients tested, the Dearborn test was given to 36 normal adults connected with the Psychopathic Hospital. Most of these were physicians, social workers and psychologists. Although this is a group test, it was given individually to many of the persons tested, both patients and normal adults.

The total number of patients tested was 77. Seven correlations were calculated by the ordinary Pearson formula. (The character of the data rendered rank difference methods impracticable.) First, the Stanford-Binet and Dearborn scores of the entire group were correlated. The r was found to be .52. The cases were then grouped according to diagnosis. It should be noted here that no patients were tested who were obviously psychotic, or whose scores would probably be lowered by their condition at the time of the test. Seventeen fell into diagnostic groups which were too small to be dealt with separately. The correlations between Stanford-Binet scores and Dearborn scores of the groups large enough to use were as follows. Using 14 cases diagnosed psychopathic personality r was .36. When the records of two of the patients in this group were omitted, because of having driven cars under the same traffic rules as those in the third and fourth mazes of the Dearborn test, the correlation of the remaining twelve patients was 67. Fourteen cases diagnosed mental retardation (subnormal intelligence) gave a correlation of —62. In this group, there were four cases of life ages respectively 15-4, 14-9, 12-11, 10-6. These 14 combined with a group of 8 cases diagnosed mental retardation plus conduct disorder, gave a correlation of —17. A group of 13 cases of conduct disorder gave a correlation of —15. When these cases were combined with 7 others diagnosed conduct disorder, plus mental retardation, the correlation found was 7.

In the group of normal adults the scores were consistently higher throughout than in any of the groups of patients. The accompanying

distribution compares the Dearborn scores of the normal adults, the cases of psychopathic personality and a group of the adult patients tested whose mental ages fall below 13 years, thirteen being considered the lower limit of average adult intelligence. Three cases, one at score 15, one at score 19 and one at score 23, refer to the 3 individuals (2 patients and 1 physician) who have driven cars in cities under the traffic rules used in the third and fourth mazes. The median Dearborn score for the normal adults is 12.5, for the cases of psychopathic personality 5.5, and for the patients with mental age below thirteen, 3.3.

Dearborn Score	0	1	2	3	4	5	6	7	8	9	10	11	12
Normal Adults	0	0	0	1	1	1	1	1	2	6	3	1	2
Psychopathic Personalities	1	0	2	2	2	0	0	2	0	0	1	0	0
Patients with Mental Ages below 13	3	2	6	6	8	3	0	0	0	0	0	0	0

Dearborn Score	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Normal Adults	2	2	2	1	0	2	0	0	0	2	3	0	0	3
Psychopathic Personalities	0	1	1	0	1	0	1	0	0	0	0	0	0	0
Patient with Mental Ages below 13	0	0	0	0	0	0	0	0	0	0	0	0	0	0

While the ability concerned in the maze tests is seen usually to go hand in hand with the more general intelligence criterion, it at times plays a more special rôle, notably below or above the general intelligence rating. Among the cases under diagnosis of mental retardation, this is so frequent as to give the expression of relationship its markedly negative value, —62. This is evidence of more disparate development of intelligence functions in this class of cases.

COMMUNICATIONS AND NOTES

INDUSTRIAL PSYCHOLOGY IN EUROPE

To the Editors of the Journal of Applied Psychology:

Sirs: I have just returned from a visit to the laboratories of industrial psychology in Europe, and inasmuch as several colleagues have suggested that a report of my observations would be of interest to the readers of the JOURNAL OF APPLIED PSYCHOLOGY, I take pleasure in submitting an informal résumé of my tour.

In all the countries visited, I found the applied psychologists interested in this field, though stressing divers points in approaching it.

In England, industrial psychology has a sort of official status, one important branch being under the auspices of the Industrial Fatigue Research Board (something like our National Research Council), which has been for several years investigating the relation between fatigue and various working conditions. Some publications of the Board are already available, a list of which may be secured upon application to the secretary at 6 John St., Adelphi, London, W. C. 2. The nature of the current investigations may be denoted by one which Miss Smith is conducting upon the conditions under which women work in laundries. I was struck with special force by the novelty of some of the methods that have been devised for measuring fatigue—not, be it said, by means of the oft-tried mechanical instruments; the attempt is made, rather, to express it as a function of the job itself.

Another organization that indicates the English interest in industrial psychology is the National Institute of Industrial Psychology, 329 High Holborn, London.

Though the operating center of these two organizations is London, the animating center is largely Cambridge. Dr. Myers has been indefatigable in his efforts to inaugurate and secure support for the work. As part of his program he called Mr. B. Muscio from Sydney to lecture at Cambridge upon industrial psychology and to conduct investigations in the laboratory.

As I shall presently show by comparison, the English approach emphasizes one point of view particularly. For example, tests do not bulk as large there as in some other centers. Nevertheless I am very optimistic regarding the results of the English approach to the problem. It has bearings beyond its ostensible aim of ameliorating conditions of work. It contitutes a long step in the direction of the too-long neglected analysis of the job, and as such it may furnish information useful in vocational selection and guidance, in increasing output, and in many other economic functions.[1]

In France I found lively interest in the psychological problems of industry, though it is scattered among several circles. A great part exists as *le taylorisme*, which has developed into a veritable cult among French engineers. Its vogue has penetrated even into conservative university circles, as evidenced by the course given at the Sorbonne by M. Charles Cestre, upon American methods of industrial management. M. Cestre introduced me to a half dozen of the leading economists

[1] See H. D. Kitson, "Scientific Method in Job Analysis," *J. of Pol. Econ.*, XXIX (June, 1921), 508-514.

and engineers who have organized for the sake of giving impetus to such movements, the *Société d'encouragement au commerce et á l'industrie*, 4 rue de Rennes, Paris.

The chief investigations emanating from orthodox psychological sources are those carried on by M. J. M. Lahy at Paris. He has written a book, "*Le System Taylor et la physiologie du travail professionel.*" He is also secretary of the 1921 conference of industrial psychology which I shall mention presently.

Further French interest in industrial problems is revealed in several physiological studies, some upon fatigue, and some upon other effects and conditions of work. There is still another angle that is represented—the pedagogical. Here the chief interest centers around vocational guidance, with considerable thought for the *mutilés* of war.

My visiting in Germany was confined exclusively to Berlin where I was conducted about chiefly by Dr. Piorkowski who, with Dr. Moede, edits the journal, *Praktische Psychologie*.[2] These investigators have an Institute with headquarters at the *Technische Hochschule* located in the Charlottenburg section of Greater Berlin. Their activities are manifold. At the *Hochschule* they administer tests to students and apprentices; in the capacity of consultants they have introduced occupational tests into several factories in Berlin; and they publish the journal.

The kinds of tests used follow largely the Muensterberg tradition of attempting to reproduce the work to be performed, in diminutive and elemental form. For example, the test for aptitude for running a machine requiring retention and reproduction of a certain pressure just made by the hand, consists of an apparatus upon which the subject can exert a pressure which is recorded, and which he afterwards reproduces as closely as possible. The two kymographic records are then compared, giving an "aptitude index." Further examples of the tests employed and the methods employed in evaluating them may be found in the pages of "Praktische Psychologie."

In order to save me time and energy in viewing all their activities, my German hosts had prepared a three-reel film showing the operation of many varieties of tests. This they presented to me in private demonstration. They informed me that they would be glad to distribute this in the United States at the price of three hundred dollars.

During the week I was in Berlin a series of intelligence tests was being administered to children just emerging from the *Mittelschule*, in order to select a certain number as entrants in a school for gifted children. A large group had already been selected on the basis of scholarship; and they were being still further sifted by means of intelligence tests. The day I observed the work about 150 girls were gathered in the grand council-chamber of the *Berliner Rathaus* for testing. The tests employed were the traditional word-building, computation, Ebbinghaus completion, and the like. The testing went very smoothly but very slowly. The children wrote on separate sheets of paper torn from booklets similar to the examination "blue-books" of American universities. For three hours that morning the children sat and took tests with only a fifteen minute intermission. The work of that morning represented only a small portion of the total undertaking, an entire week being required for testing all the children. The work of scoring papers must have been mountainous. Some of our Ameri-

[2] See H. C. Link, "A New Journal of Practical Psychology," *J. of App. Psychol.*, V (March, 1921), 85-88.

can psychologists who are specializing in forms of intelligence tests and in devices for increasing efficiency in their administration may find a good market for their wares in Berlin.

At Geneva, I found the *Institut J. J. Rousseau* likewise interested in industrial psychology. This institution, founded in 1912 under the direction of Professor Claparéde, is devoted chiefly to the application of psychology to educational problems. But in the light of recent movements in vocational education, the Institute feels itself bound to direct a moiety of its energies into the industrial field. As early as 1916 a rather pretentious attempt was made to formulate a series of tests for telephone operators. The results were published in the *Archivs de Psychologie* and furnished a pattern for several similar investigations in Europe. Occupations that the Institute has since approached analytically in various Swiss factories and technical schools are watchmaking, various operations of the machinist's trade, and several occupations for women. The report of this latter group of studies, M. Bovet, Director of the Institute, assured me, is in fair way to publication.

A considerable amount of attention is being paid on the continent to vocational guidance, or "*orientation professionelle*," as it is called. I found two books upon the subject in French, and two journals—one in French, and one in Spanish. Germany has a national association of vocational guidance of which Dr. Moede is president. And all Europe has inaugurated a series of annual conferences. The first one was held last year at Geneva. The second (*Deuxième Conférence Internationale de Psychotechnique Appliquée à l'orientation professionelle et à l'organization du travail*), will take place at Barcelona September 28-30, 1921. The program is divided into three sections: Section 1, (a) Vocational guidance; (b) motion study; (c) occupational fatigue. Section 2, Psychographic methods (especially for schools). Section 3, Tests. The conference will be entertained by the personnel of *L'Institut d'orientation professionelle* of Barcelona. I should like to have stayed over for this conference and have brought back a complete report of it but I had not made arrangements to miss the opening weeks of the university.[3]

From observations such as those I made, one may conclude that the interest in industrial psychology is just as keen in England, France, Germany and Switzerland, as it is in the United States. It is very intense, and is animated by high moral and patriotic motives of improving conditions of labor, raising the efficiency of production, guiding young people (especially those maimed in the war) into appropriate occupations, and educating them with effectiveness and dispatch. In each of the countries emphasis is placed upon one or another of these phases. In England, perhaps it is chiefly directed toward the amelioration of working conditions; in France, toward raising general technical efficiency; in Switzerland toward vocational guidance, or at least, vocational education.

Intense though the interest is, it is not as extensive, however, as that in the United States. (There were only about twenty in attendance at the last meeting of applied psychologists at Geneva.) One receives the impression that European psychologists look with considerable admiration upon the achievements of American psychologists in

[3] The editors are pleased to announce that through Mr. Kitson's efforts they hope to present the outstanding features of the conference in a later issue.

this field, particularly upon the personnel activities of the army. (I think I am betraying no professional or military secrets in remarking that there has been serious talk of adopting our personnel methods in the Swiss army.)

The European investigators are generally acquainted with our books upon the subject, but not with our periodical literature. Nowhere did I find in use the recent forms of our twenty-seven varieties of intelligence tests. Nowhere did I find test-booklets, "cross-out" elements, scoring stencils, simplified statistical methods and other aids to efficiency so common in the States. I suggest, therefore, that any American mental testers who are anxious to spread the use of their tests may profitably place on their mailing-lists the following European names:

Dr. C. S. Myers, Psychological Laboratory, Cambridge University.
Mr. B. Muscio, Psychological Laboratory, Cambridge University.
Mr. G. H. Miles, 329 High Holborn, London, W. C.
Mr. D. R. Wilson, 6 John St., Adelphi, London, W. C. 2.
Mr. Eric Farmer, 6 John St., Adelphi, London, W. C. 2.
Miss May Smith, 6 John St., Adelphi, London, W. C. 2.
Dr. W. Moede, Berlin, W. 30, Luitpoldstrasse, 14.
Dr. C. Piorkowski, W. 30, Luitpoldstrasse, 14.
Professor Ed. Claparéde, Psychological Laboratory, University of Geneva.
M. Heinis, Psychological Laboratory, University of Geneva.
M. Pierre Bovet, Director Institute J. J. Rousseau, Geneva.
Madame Chavennes, Institute J. J. Rousseau, Geneva.
Mlle. Chavennes, Institute J. J. Rousseau, Geneva.
M. J. M. Lahy, 12 rue Linné, Paris.

The trip showed me that the problems which confront industrial psychologists are similar, whether in Europe or America, and that when approached from the scientific point of view, they lead over quite similar pathways. It also showed me that in spite of the mere three thousand miles that separate our coasts there is regrettable lack of acquaintance between applied psychologists in Europe and America. And that in order for a maximum of progress to be made by both groups there must be deliberate and constant efforts towards faciliating a *rapprochement*.

Professor of Psychology, Indiana University,
Bloomington, Indiana.

HARRY DEXTER KITSON

THE NATIONAL INSTITUTE OF INDUSTRIAL PSYCHOLOGY OF THE UNITED KINGDOM

For some time the more progressive firms in this country have been interested in the application of psychological and physiological principles to industry and commerce. The Industrial Fatigue Research Board which was established in 1915 has already done much to show the possibilities in this field. Its activities are however necessarily confined to broad surveys and investigations of common problems and it was felt by many firms interested, that much of the value of its work would be wasted unless assistance could be obtained in the application of the principles to the special needs of individual firms. This work is obviously beyond the scope of a Board supported by public funds and in order to meet this need a National Institute of Industrial Psychology has been formed. This was incorporated under the Com-

panies Act on February 11, 1921, as an Association for Scientific Research. It has the support of a number of well-known business men and in addition has a Scientific Advisory Committee consisting of the heads of psychological, physiological and educational departments of Universities throughout the Kingdom who are interested in the practical application of the sciences of Psychology and Physiology. Private individuals have come forward with generous financial support, and in addition the Carnegie United Kingdom Trust is giving substantial help. Where investigations are carried out which are likely to result in an immediate advantage to a firm, fees are charged which cover the salaries of the investigators and include overhead charges of the central body.

The Institute is under the direction of Dr. C. S. Myers, M. D., Sc. D., F. R. S., and by arrangement with the Industrial Fatigue Research Board the services of their investigators may be utilized when required to supplement the work of the Institute's investigators.

On the industrial side of the Institute's activities investigations are now being conducted in factories concerned in the work of presses, box and can making, soldering, chocolate dipping and packing: interesting problems concerning layout, size of material, rest pauses, selection tests, seating accommodation and ventilation have arisen and are being examined.

On the commercial side the formation and trying out of a suitable series of selection tests for the clerical staff of large firms is being undertaken, which promises to prove most useful.

When more detailed information is obtained as a result of these and other investigations, the Institute will be in possession of material invaluable for Vocational Guidance. By cooperation with a number of leading education authorities definite and scientific lines of approach will be formed between the school and business life so that a boy or girl, on leaving school, will have the opportunity of obtaining more exact and scientific guidance than has been possible in the past. In this way it is hoped that much of the waste of national energy which takes place in the intermediate stage by haphazard drift and false starts will be avoided.

The Institute is therefore making a more widespread appeal to all those who are interested in this work. That it has been possible to advance so far in a time of such severe financial depression is a hopeful promise of greater activity in brighter times. Meanwhile active propaganda work is being carried on, both amongst employers and employees, and lectures dealing with the subject are being arranged. The results of the work and a full survey of the work of similar character carried out by other similar bodies abroad will be published by the Institute in the form of a quarterly journal.

The Secretary of the Institute is Dr. G. H. Miles and the head offices are at 329 High Holborn, London, W. C. 1.

We are indebted for the above account of the work of the National Institute of Industrial Psychology of the United Kingdom to the Director, Dr. Charles S. Myers.

BOOK REVIEWS

KEMPF, EDWARD J. *Psychopathology*. St. Louis, C. V. Mosby Co., 1920. pp. 762.

"It is comparatively simple and very easy to be an extremist in anything, but it requires eternal care, sound common sense, and no little patience and endurance to maintain a progressively refining, healthfully constructive attitude toward the fundamental needs and pleasures of human nature." When the above statement appears as early as the third page of this recent contribution to psychopathological literature, the reader is inclined to assume a rather comfortable postural tonus and approach the volume with the expectation that perhaps at last has been found the long anticipated volume which is to safely point a course between the extremes of too rigid conservatism and too liberal psychoanalytic deductions. Surely a writer who voices the need of "eternal care" and "sound common sense" will not fail to heed his own warning. For the reader expecting a conservative interpretation of human adjustment, there is increasing disappointment. He will find here outlined, no medium course, but still another extreme contribution to psychopathological literature, which he must appraise and, after selection and rejection, weave into his own hypotheses.

Of one thing there can be no doubt. Kempf himself is sincerely convinced of his theory and is most eager to proselyte. In fact, his enthusiasm is so contagious that there is danger of uncritically accepting statements for which reason forces, if not rejection, at least serious question. In actual presentation of data, Kempf seems so anxious to get his view point before the public, that he takes no time to refine. As a result there is much to be criticized in his style. It is neither consistently popular nor consistently scientific. One finds tedious repetition, unnecessarily long and involved sentences, an occasional incomplete sentence and illustrations ad infinitum. It is to be regretted that "Psychopathology" is not presented with the fineness of expression of "The Yellow Jacket." Unkindly criticism might characterize Kempf's present style as he does his patients. "The large number of unique words which he used to express a simple thought was unusually impressive." (Page 436.) Again, "he used numerous polysyllabic words with little regard to their meaning." (Page 530.) It would indeed take a clever reader to grasp the meanings of "goodness" and "happiness" from a study of the wordy definitions on page 118. However, it is his coarseness of expression which can least be forgiven. Kempf himself calls attention (page 3) to the biological destructiveness of the tendencies to vulgarity and to prudishness. Certainly the way to overcome "prejudice, prudery, or mock finery" is not by actual coarseness of expression. The latter is quite unnecessary, however "delicate" the situation discussed.

To understand Kempf's dynamic, autonomic mechanism of the personality, one must appreciate first of all the tremendous emphasis placed upon the autonomic system. He assumes that "the different segments of the autonomic apparatus are stimulated to assume different types of postural tensions and activities, which give rise to an affective nervous stream, which, in turn, coordinates the projicient apparatus and compels it to act so as to expose the receptors of the

organism so that they will acquire certain types of stimuli and avoid others. The stimuli must be acquired in order to avoid prolonged unrest and distress, which may become decidedly malnutritional in their influence, must have the capacity to counter-stimulate the autonomic segment so that it will resume a state of comfortable tonus." (Pages 9-10.) "Emotions and sentiments are cravings that have their origin in the tensions and movements of different autonomic (visceral) segments." (Page 6.) If emotions have a peripheral origin in the characteristic variation of the postural tensions or visceral segments, it is clear that at a given time, the individual, although he may be dominated by a distinct craving, nevertheless is never possessed by a single affective craving. Gradually the autonomic segments become conditioned to seek only such stimuli as are approved by the social group. Will becomes the wish for the esteem of the group. With the integration into a unity of these segmental cravings, together with one's strivings to win social esteem, there develops a personality, the Ego. Normally, of course, only appropriate stimuli would satisfy segmental cravings and so neutralize tension, but unfortunately, it is quite possible for the organism to become conditioned to demand less appropriate stimuli and thereby we have within a single personality the basis for possible conflict between segmental cravings and the striving for social esteem. Out of such conflict arise the mechanisms of suppression, repression, dissociation, and sublimation. In general, then, to Kempf behavior becomes largely a problem of supplying the maximum gratification with the minimum expenditure of energy for segmental cravings, at the same time preventing gratification at the expense of social esteem. Eccentric behavior becomes a compensation for perverse cravings. A segment denied reality must have a substitute. Delusions, hallucinations, and dreams, however bizarre, are but means of gratifying cravings which cannot be gratified by reality.

In appraising Kempf's theory two questions arise: First, what of the validity of the physiological data, the very foundation on which the psychological structure rests? Second, are we justified in considering the affect as purely autonomic in origin? For the physiological data, Kempf finds justification for his point of view largely in the work of Sherrington, Langelaan and Cannon. Even a casual reader might question the sweepingness of certain of Kempf's deductions from the findings of these men. Adequate criticism belongs to the neurologist or physiologist rather than psychologist. So far as the second question is concerned, does not Cannon himself raise the question whether organic changes are sufficiently differentiated in various emotions to account for the affective differences?

As regards Kempf's interpretation of behavior there is indeed much that is illuminating, particularly in his emphasis upon the importance of the Ego's attitude toward the cravings as determining whether his adjustment shall be one of submission, elimination, simulation or sublimation. Certainly such an explanation renders intelligible much that might seem utterly unintelligible, particularly in pathological behavior. Whereas by and large such mechanisms may be acceptable, there are times when the interpretation invites reductio ad absurdum. One may grant that Darwin may have had an inferiority complex and yet question whether it be necessary to assume that his use of a high stool in studying was to elevate himself above his fellows, thereby compensating for his feeling of inferiority. One may grant that a patient attempts to compensate for his impotence through the construction of a perpetual motion machine, and yet hesitate to see in invention in

general a compensation for the wastes of masturbation. One familiar with the number of objects on a psychopathic ward which would be available for swallowing, can hardly fail to feel the overemphasis on the swallowing of the ticking watch by the woman patient in simulation of pregnancy. Certainly an interpretive analysis of cases should be substituted for the older description of symptoms. Kempf's "Psychopathology" has been of no little value if it stimulates interest in case analyses. Whether one must ultimately approach all such analyses from the autonomic slant is another question.

As would be expected Kempf makes generous use of the psychoanalytical stock of symbols. In his case studies, their use is what might be expected from any thorough going analyst. However, his interpretation of the symbolic significance of various works of art, suggests that he can see only what he is looking for. For example, he cites Rodin's "Caryatid"—a woman's figure crouched beneath the weight of a burden resting on her shoulder. At once, Kempf sees in the bowl only the psychoanalytic symbol for the uterus, the shoulder is a convenient substitute, and the statue becomes symbolic of the "burdensome uterus and longing for maternity." It is interesting to question to what extent one is justified in reading an interpretation into another's creation. One might also raise the question whether the contemplation of this statue would afford relief from repressions of the type indicated. So far as the psychology of art is concerned, stone obviously from the inherent qualities of the substance itself is adapted to a representation of the type or symbolic rather than the individual or pictorial. But why must all that is symbolic be symbolic of libido at a crude organic level?

Certain it is that no one can contemplate Kempf's interpretations of the mechanisms of human adjustment without finding much that is illuminating and stimulating. Accepted in toto, however, it is legitimate to ask whether such a point of view is conducive to the development of that much to be desired "progressively refining, healthfully constructive attitude toward the fundamental needs and pleasures of human nature." Is an interpretation of life which sees the great driving forces of human behavior as purely organic, conducive to the development of a civilization that will survive? Regardless of its acceptability or rejection on purely emotional grounds, is it not clear that such an ideal would be biologically and ethically a regression? The ultimate test of course is pragmatic. Kempf freely criticizes the "biological castration influence of present American educational and social tendencies" but does not adequately outline a program for their reconstruction. Until he does, it is left to the imagination of the reader to reconstruct a social order in accordance with an ideal of an organic summum bonum. Such an attempt at reconstruction makes particularly clear the social and ethical implications of an autonomic ideal and protects against too uncritcal acceptance.

The following books and pamphlets have been received:

James Drever.—*The Psychology of Industry*. E. P. Dutton Co., New York, 1921. 148 pp.

—— ——. *The Psychology of Everyday Life*. E. P. Dutton Co., New York, 1920. 164 pp.

Henry R. Evans. *Educational Boards and Foundations*, 1918-20. Washington Govt. Printing Office, 1921. 11 pp.

Julia Wade Abbot. *Kindergarten Education*, 1918-20. Washington Govt. Printing Office, 1921. 12 pp.

Dr. A. A. ROBACK. *Roback Mentality Tests for Superior Adults.* Cambridge, Mass. 14 pp.

JOHN J. B. MORGAN. *Morgan's Mental Test.* Psychopathic Hospital State University of Iowa. 8 pp.

HELEN T. WOOLEY, Ph.D. and HORNELL HART. *Feeble-Minded Ex-School Children.* Studies from the Helen S. Trounstine Foundation, 25 East Ninth Street, Cincinnati, Ohio. April 1, 1921. 27 pp. Price 50 cents; 4 or more copies, 25 cents each; in set with 6 other studies $1.00.

H. R. BONNER. *Statistics of Universities, Colleges and Professional Schools,* 1917-18. Washington Gov't Printing Office, 1921. 223 pp.

WM. R. HOOD. *State Laws Relating to Education.* Washington Gov't. Printing Office, 1921. 231 pp.

CHAS. S. MYERS. *Mind and Work.* C. P. Putnam's Sons, New York and London. The Knickerbocker Press, 1921. 175 pp.

HARRY DEXTER KITSON, PH.D. *The Mind of the Buyer.* The Macmillan Co., New York, 1921. 206 p.

ALEXANDER INGLIS. *Inglis Intelligence Quotient Values.* World Book Co., Yonkers-on-Hudson, New York. 16 pp.

Directory of American Psychological Periodicals